

Non-respons en rotatie in het Bedrijven-Informatienet

Kwantitatieve en kwalitatieve analyse van de effecten

Dr. Hans C.J. Vrolijk
Drs. Geerte C. Cotteleer

Projectcode 63002

April 2005

Rapport 1.05.01

LEI, Den Haag

Het LEI beweegt zich op een breed terrein van onderzoek dat in diverse domeinen kan worden opgedeeld. Dit rapport valt binnen het domein:

- ☒ Wettelijke en dienstverlenende taken
- ☐ Bedrijfsontwikkeling en concurrentiepositie
- ☐ Natuurlijke hulpbronnen en milieu
- ☐ Ruimte en Economie
- ☐ Ketens
- ☐ Beleid
- ☐ Gamma, instituties, mens en beleving
- ☐ Modellen en Data

Non-respons en rotatie in het Bedrijven-Informatienet; Kwantitatieve en kwalitatieve analyse van de effecten

Vrolijk, H.C.J. en G.C. Cotteleer

Den Haag, LEI, 2005

Rapport 1.05.01; ISBN 90-5242-983-9; Prijs €12,25 (inclusief 6% BTW)

58 p., fig., tab., bijl.

In het Bedrijven-Informatienet van het LEI (het Informatienet) wordt een gedetailleerde administratie bijgehouden van 1.500 land- en tuinbouwbedrijven. Naast financieel-economische gegevens worden ook technisch-economische gegevens van deze bedrijven vastgelegd. De kwaliteit van de resulterende gegevens is onder andere afhankelijk van de bedrijven die in de steekproef worden opgenomen. Jaarlijks wordt een deel van de bedrijven vervangen. Dit rapport analyseert twee belangrijke aspecten van de jaarlijkse aanvulling van de steekproef, namelijk de non-respons en de rotatie.

In the Dutch Farm Accountancy Data Network, accounting information of 1,500 horticultural and agricultural farms is collected. Not only financial economic information, but also technical economic information is collected. The quality of the estimates based on this sample depends on the farms that are included in the sample. This report analyses two important aspects of the yearly completion of the sample, non-response and rotation.

Bestellingen:

Telefoon: 070-3358330

Telefax: 070-3615624

E-mail: publicatie.lei@wur.nl

Informatie:

Telefoon: 070-3358330

Telefax: 070-3615624

E-mail: informatie.lei@wur.nl

© LEI, 2005

Vermenigvuldiging of overname van gegevens:

- ☒ toegestaan mits met duidelijke bronvermelding
- ☐ niet toegestaan



Op al onze onderzoeksopdrachten zijn de Algemene Voorwaarden van de Dienst Landbouwkundig Onderzoek (DLO-NL) van toepassing. Deze zijn gedeponeerd bij de Kamer van Koophandel Midden-Gelderland te Arnhem.

Inhoud

	Blz.
Woord vooraf	7
Samenvatting	9
Summary	13
1. Deel I: Non-respons - Inleiding en doelstelling	17
1.1 Inleiding	17
1.2 Doelstelling	17
1.3 Fouten in steekproefonderzoek	17
1.4 Non-responsfouten	18
1.5 Steekproefffouten	19
1.6 Opzet van onderzoek	20
2. Analyse non-respons	21
2.1 Kwalitatieve analyse redenen non-respons	21
2.2 Kwantitatieve analyse non-respons	23
2.3 Kwantitatieve analyse non-responsbias	25
2.4 Ongeschikt zijn van bedrijven	27
3. Representativiteit van het Informatienet	29
3.1 Definitie van representativiteit	29
3.2 Representativiteit van het Informatienet	30
4. Conclusies en aanbevelingen non-respons	33
5. Deel II: Rotatiemethodieken - Inleiding en doelstelling	35
5.1 Inleiding	35
5.2 Doelstelling	36
5.3 Opzet van onderzoek	36
6. Rotatiemethodieken en randvoorwaarden	37
6.1 Inleiding	37
6.2 Huidige opzet Informatienet	37
6.3 Randvoorwaarden	37

	Blz.
6.4 Rotatiemethodieken	38
6.4.1 Minimale rotatie	38
6.4.2 Huidige rotatiemethode	39
6.4.3 Halve rotatie	39
6.4.4 Praktische rotatie	39
6.4.5 Volledige rotatie	40
7. Voor- en nadelen van rotatie	41
7.1 Inleiding	41
7.2 Bruikbaarheid voor het onderzoek	41
7.3 Administratiekosten	43
7.4 Leereffect	44
7.5 Non-respons bij aanvang	45
7.6 Uitval deelnemende bedrijven	46
7.7 Structuurveranderingen landbouw	46
7.8 Kwaliteit gegevensverzameling	47
7.9 Praktische mogelijkheden van de gegevensverzameling	47
7.10 Conclusies	50
8. Conclusies en aanbevelingen rotatie onderzoek	51
Literatuur	53
Bijlagen	56
1. Samenhang Informatienet-deelname en prestaties bedrijf	56
2. Berekening benodigde capaciteit per methode	58

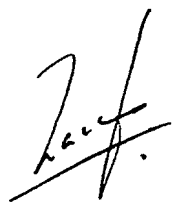
Woord vooraf

In het Bedrijven-Informatienet wordt een gedetailleerde administratie bijgehouden van ruim 1.500 land- en tuinbouwbedrijven. Naast financieel-economische gegevens worden ook technisch-economische gegevens van deze bedrijven vastgelegd. De kwaliteit van de resulterende schattingen is onder andere afhankelijk van de bedrijven die in de steekproef worden opgenomen.

Jaarlijks wordt een deel van de bedrijven vervangen. Naar aanleiding van het steekproefplan wordt de gewenste samenstelling van de steekproef vastgesteld. De meest recente Landbouwtelling wordt gebruikt om de bedrijven door middel van een gestratificeerde steekproef te selecteren volgens dit steekproefplan. De bedrijfskeuze richt zich op het aanvullen van de steekproef.

Dit rapport analyseert twee belangrijke aspecten van de jaarlijkse aanvulling van de steekproef, namelijk de rotatie en de effecten van non-respons. De op dit moment gehanteerde rotatiemethode wordt geëvalueerd. Daarnaast worden andere mogelijke rotatiemethoden onderzocht. Ten aanzien van de non-respons wordt geanalyseerd in hoeverre bedrijven die mee willen werken afwijken van de bedrijven die weigeren en wat de effecten hiervan zijn op de resultaten. Dit rapport geeft aanbevelingen over de herziening van de rotatiemethodiek en het minimaliseren van de effecten van non-respons.

Dit rapport is opgesteld door Geerte Cotteleer en Hans Vrolijk. De onderzoekers willen Bernard Douma en Hennie van Welzen bedanken voor hun bijdrage aan de discussie omtrent rotatie. Karel Lodder wordt bedankt voor zijn bijdrage aan de opzet van het onderzoek.



Prof.dr.ir. L.C. Zachariasse
Algemeen directeur LEI B.V.

Samenvatting

In het Bedrijven-Informatienet van het LEI (het Informatienet) wordt een gedetailleerde administratie bijgehouden van 1.500 land- en tuinbouwbedrijven. Naast financieel-economische gegevens worden ook technisch-economische gegevens van deze bedrijven vastgelegd. De kwaliteit van de resulterende gegevens is onder andere afhankelijk van de bedrijven die in de steekproef worden opgenomen. Jaarlijks wordt een deel van de bedrijven vervangen. Dit rapport analyseert twee belangrijke aspecten van de jaarlijkse aanvulling van de steekproef, namelijk de non-respons en de rotatie.

Non-respons

Jaarlijks worden bedrijven benaderd om deel te nemen aan het Informatienet. Een deel van deze bedrijven zal echter weigeren deel te nemen. Dit hoeft op zich geen probleem te zijn als deze bedrijven niet afwijken van andere bedrijven. Als de bedrijven die weigeren systematisch bepaalde kenmerken hebben zullen deze kenmerken ondervertegenwoordigd zijn in de steekproef. Als bijvoorbeeld oude agrariërs minder wensen mee te werken leidt dit ertoe dat de leeftijdsverdeling van de deelnemers in het Informatienet niet representatief is voor de leeftijd in de populatie (in hoeverre dit verschil van invloed is op de uitkomsten van een specifiek onderzoek hangt af van het onderzoeksdoel en in hoeverre de daarvoor relevante kenmerken samenhangen met de leeftijd).

Uit de analyse van de non-responsbias blijkt dat er een licht verschil bestaat tussen de responderende en niet-responderende bedrijven. Deze verschillen zijn echter niet significant. Opmerkelijk is wel dat de bedrijven die door de Technisch Administratief Medewerker (TAM) als ongeschikt worden beschouwd voor deelname in het informatienet, zowel qua structuur als qua resultaten significant afwijken van de andere twee groepen. In samenhang met de constatering dat het percentage ongeschikte bedrijven sterk uiteenloopt tussen verschillende wervers, leidt dit tot de aanbeveling de criteria voor het al dan niet geschikt zijn van bedrijven duidelijker te formuleren en te communiceren.

Uit de analyse van de vragen omtrent de houding en mening van de benaderde boeren blijkt dat met name het vertrouwen in de overheid en de mate waarin de boer op de hoogte is van het Informatienet bepalen of een boer al dan niet deelneemt. De eerste factor 'vertrouwen in de overheid' is moeilijk door het LEI te beïnvloeden. Op de tweede factor 'op de hoogte van het Informatienet' heeft het LEI deze invloed wel. Middels een betere communicatie via bijvoorbeeld vakbladen kan ervoor gezorgd worden dat potentiële deelnemers aan het Informatienet beter op de hoogte worden gesteld van de manier waarop het Informatienet wordt gebruikt.

Uit de analyses omtrent de representativiteit van het Informatienet blijkt dat het Informatienet in grote mate representatief is voor de belangrijke stratificatievariabele bedrijfsomvang (ege). Ook voor oppervlakte resulteert het Informatienet in goede schattingen. Voor variabelen die verder afstaan van de variabelen die bij de opzet van het

Informatienet zijn gebruikt, moet in specifiek onderzoek worden nagegaan of de opgenomen bedrijven inderdaad representatief zijn.

Uit de gegevens blijkt verder dat er grote verschillen bestaan tussen de TAM's die bedrijven werven (vanwege de privacy zijn deze resultaten niet opgenomen in dit rapport). De verschillen zitten zowel in de responspercentages als het percentage bedrijven dat als ongeschikt wordt aangemerkt. Ten dele kunnen deze verschillen worden verklaard uit de bedrijfstypen waarmee een TAM vooral werkt. Er is echter wel degelijk een persoonseffect. Bij toekomstige wervingen verdient het aanbeveling vooral die TAM's te laten werven die er lol aan hebben, de betekenis van het Informatienet duidelijk kunnen communiceren en zelf overtuigd zijn van het nut van het Informatienet.

Uit de analyse van het niet geschikt zijn van bedrijven blijkt dat elk jaar een groot deel van de bedrijven als ongeschikt voor deelname aan het Informatienet wordt bestempeld. Ten dele wordt dit veroorzaakt door praktische problemen zoals het herhaaldelijk onbereikbaar zijn van een bedrijf. Daarnaast geldt dat veel bedrijven ondertussen zijn gestopt. Dit maakt de bedrijven inderdaad ongeschikt voor deelname aan het Informatienet, maar roept wel de vraag op hoeveel soortgelijke bedrijven er in de landbouwtelling voorkomen. Er zijn echter ook meer subjectieve aspecten waarom een bedrijf als ongeschikt wordt bestempeld. Nevenactiviteiten of het ondergeschikt zijn van de agrarische activiteiten leiden er vaak toe dat een bedrijf niet wordt opgenomen. Systematisch (kleine) bedrijven als niet ter zake doende af te wijzen, leidt tot een slechte vertegenwoordiging van deze bedrijven. Het is te overwegen deze agrarische activiteiten wel in het Informatienet op te nemen, of in de Europese wetgeving moet de 16-ege-grens worden herzien.

Rotatie

Het Informatienet heeft de vorm van een roterend panel. Een panel is een combinatie van een cross-sectie en een tijdreeks. Een cross-sectie is een dwarsdoorsnede in de tijd. Dit betekent dat er gekeken wordt naar gegevens over een aantal verschillende variabelen, objecten of entiteiten op één tijdstip of in een (vaste) periode. Een tijdreeks bevat gegevens over één variabele op een aantal tijdstippen. Bij een panel worden gegevens van meerdere variabelen op een aantal verschillende tijdstippen beschouwd. In het Informatienet wordt jaarlijks een deel van de bedrijven vervangen. Het Informatienet is hiermee een voorbeeld van een roterend panelbestand.

Gebruikers van het Informatienet geven echter al geruime tijd aan dat zij graag zouden werken met langere tijdreeksen. Om deze reden wordt de rotatie van het Informatienet in dit onderzoek tegen het licht gehouden, en wordt nagegaan of de rotatieopzet aan herziening toe is.

In dit onderzoek worden verschillende rotatiemethoden met elkaar vergeleken. De eerste methode is minimale rotatie. Dat wil zeggen zo min mogelijk bedrijven vervangen onder de gestelde randvoorwaarden. Halve rotatie wil zeggen dat 50% van de deelnemers aan het Informatienet jaarlijks in de steekproef wordt gehouden en 50% jaarlijks vervangen (indien mogelijk wordt dezelfde 50% jaarlijks vastgehouden). Huidige en praktische rotatie lijken sterk op elkaar. Beiden komen neer op een rotatiepercentage tussen de 20 en 33%. Volledige rotatie wil zeggen dat alle bedrijven jaarlijks worden vervangen.

In het rapport zijn de rotatiemethoden geëvalueerd aan de hand van een verzameling

theoretische en praktische criteria. In tabel 1 zijn alle rotatiemethoden aan de hand van de verschillende criteria beoordeeld.

Tabel 1 Voor- en nadelen van alternatieve rotatiemethoden

Rotatiemethode Criterium	Minimaal	Huidig	Praktisch	Half	Volledig
Bruikbaarheid voor onderzoek op de lange termijn (paneldata- en tijdreeksanalyse)	++	+/-	+/-	-	--
Bruikbaarheid voor onderzoek van éénjarige gemiddelden	+/-	+	+	+	++
Bruikbaarheid voor meerjarige gemiddelden	+	+	+	+/-	--
Bruikbaarheid bij het inschatten van veranderingen	++	+	+	+/-	--
Benodigde capaciteit/administratiekosten	++	+/-	+/-	-	--
Leereffect: Bedrijven zijn niet meer representatief doordat zij van deelname leren	-	+/-	+/-	+/-	++
Uitval deelnemende bedrijven	--	-	-	+/-	+/-
Volgen van Structuurveranderingen in de landbouw	-	+/-	+/-	+/-	++
Non-respons bij werving	--	+/-	+/-	+/-	++
Kwaliteit van de gegevensverzameling	++	+/-	+/-	+/-	--
Praktische mogelijkheden gegevensverzameling	++	+/-	+/-	+	--

In de afweging tussen de verschillende methoden blijkt dat de minimale rotatie het beste gewaardeerd wordt als de beoordeling plaatsvindt aan de hand van de bovenstaande criteria. Hierbij maakt het geen verschil of het leereffect en de uitval van deelnemende bedrijven wel of niet worden meegenomen (dit zijn geen eenduidige criteria). Minimale rotatie kan worden gezien als een natuurlijk roterend panel, waarin bedrijven worden vervangen indien zij afvallen of omdat zij niet meer binnen de grenzen vallen van de strata zoals aangegeven in het steekproefplan. Het rapport sluit af met de aanbeveling over te gaan op minimale rotatie.

Summary

Non response and rotation in the Dutch Farm Accountancy Data Network; Quantitative and qualitative analyses of effects

In the Dutch Farm Accountancy Data Network (FADN), accounting information of 1,500 horticultural and agricultural farms is collected. Not only financial economic information, but also technical economic information is collected. The quality of the estimates based on this sample depends on the farms that are included in the sample. This report analyses two important aspects of the yearly completion of the sample, non-response and rotation.

Non response

Every year, farms are asked to participate in the system. Some of these farms will refuse to participate. These refusals do not cause problems if these farms do not differ from farms that participate. In case farms that refuse to participate systematically differ from the participating farms this could result in a bias. If for example older farmers are less inclined to participate, this results in a different age distribution in the sample compared to the population. The representativeness of the system on this variable will become questionable (whether this is a problem or not depends on the research goals and the extent to which the important variables correlate with age).

The analysis of the non-response, presented in this report, show a slight difference between participating and non-participating farms. However, these differences are not significant. A remarkable and significant difference is found between the unsuitable and suitable farms. The farms that are considered unsuitable to participate in the system are smaller than the suitable farms. This conclusion, together with the observation that the percentage of farms considered as unsuitable differs strongly between the recruiters, leads to the recommendation to make the criteria for unsuitable farms more explicit and to improve the communication about these criteria.

Analysis about the attitude of farmers that are asked to participate show that 'trust in the government' and the extent to which the farmer has knowledge about the FADN system are important factors in predicting the participation of an individual farmer. The trust in the government is a factor that can hardly be influenced by LEI. LEI can influence the extent to which farmers have knowledge about the FADN. Due to a better communication with the agricultural sector, the knowledge of the FADN system can be improved.

The report shows that the FADN system is to a large extent representative for the agricultural population. This is especially true for the variables that are used as stratification variables, such as the size of the farm (expressed in European size units). The system also results in good estimates for the utilised agricultural area. In specific research projects the representativity of the system should be checked for the important research variables. This is especially true when these variables are less linked to the stratification variables.

The data show large differences between the accountants of LEI who are involved in the recruitment of farms (these results are not shown in the report because of the privacy of employees). Differences are found in the response rates as well as the percentage of farms

that are considered as unsuitable. Partly these differences can be explained by the types of farming the accountant is involved in (response rate in dairy farming is for example much higher than in intensive livestock farming), but there is a significant person effect. In future years, it is therefore advisable to assign recruitment tasks to employees who enjoy doing these tasks, who can clearly communicate the importance and role of the FADN system and who are convinced about the usefulness of the system.

The analyses of farms that are considered to be unsuitable show that a large percentage of farms are considered unsuitable every year. Partly this can be explained by practical reasons such as not being able to contact the farmer. Furthermore, a significant number of farms is not in the farming business anymore. This indeed makes the farm unsuitable for participation in the system, but it raises the question to which extent farms in the census are not active anymore. There are also more subjective reasons why farms are considered to be unsuitable. Sideline activities, or the minor importance of the agricultural activity in the enterprise often lead to not including the farm in the sample. Systematically excluding these (smaller) agricultural activities from the sample results in a bad representation of these activities in the sample. It is worth considering to include these activities in the FADN or to increase the lower threshold of 16 esu, if 16 esu does not represent a fully functioning farm anymore.

Rotation

The Dutch FADN system is a rotating panel. A panel is a combination of cross section and time series. A cross section monitors a set of variables on a specific moment in time. A time series consider the same object at different moments in time. In a panel, data are collected on different units at different moments in time. In the FADN system part of the observation units, the farms, are replaced by newly recruited farms. This implies that the FADN system is a rotated, unbalanced, panel.

Several users of FADN data indicate that they would prefer longer time series on individual farms. In this report the rotation method is evaluated and recommendations are given to change the method.

Several methods are compared in this report. The first method is minimal rotation. This implies that as few as possible farms are replaced. Half rotation implies that 50% of the farms remain in the sample for a long period of time and the other 50% of the farms are replaced every year. The current rotation method and the practical rotation method are very similar. Both methods would imply the rotation of 20 to 33% of the farms. Full rotation means that all farms are replaced every year.

In this report, the rotation methods are evaluated based on a set of theoretical and practical criteria. In table 1 the evaluation of all methods on all criteria is summarised.

Table 1 Advantages and disadvantages of different rotation methods

Rotation method Criterion	Minimal	Half	Current	Practical	Full
Possibilities for research with a long term focus (panel data- and time series analyses)	++	-	+/-	+/-	--
Possibilities for research on one year averages	+/-	+	+	+	++
Possibilities for multi year averages	+	+/-	+	+	--
Utility for estimating changes	++	+/-	+	+	--
Required capacity / administration costs	++	-	+/-	+/-	--
Learning effect	-	+/-	+/-	+/-	++
Attrition	--	+/-	-	-	+/-
Monitoring structural changes in agriculture	-	+/-	+/-	+/-	++
Non response at recruitment stage	--	+/-	+/-	+/-	++
Quality of data collection	++	+/-	+/-	+/-	--
Practical opportunities for data collection	++	+	+/-	+/-	--

In the trade-off of these methods, the minimal rotation gets the highest score. This conclusion is regardless of the impact of the learning effect and attrition of farms (these criteria are not conclusive). Minimal rotation can be regarded as a natural evolving sample, in which farms are replaced if they stop participating, end their business or are not longer part of the field of observation as specified in the selection plan. The report ends with the recommendation to introduce minimal rotation.

1. Deel I: Non-respons - Inleiding en doelstelling

1.1 Inleiding

In het Bedrijven-Informatienet van het LEI (het Informatienet) wordt een gedetailleerde administratie bijgehouden van 1.500 land- en tuinbouwbedrijven. Op basis van deze bedrijven worden uitspraken gedaan over de hele populatie of over delen van deze populatie. Voorwaarde voor valide uitspraken is dat de bedrijven in het Informatienet een goede afspiegeling vormen van de bedrijven in de populatie ten aanzien van de kenmerken die in het onderzoek relevant zijn. Deze representativiteit wordt nagestreefd door gebruik te maken van een gestratificeerde steekproef. De populatie wordt in een aantal groepen ingedeeld en uit elk van de groepen worden bedrijven geselecteerd om deel te nemen in de steekproef. Een deel van de benaderde bedrijven zal echter niet geschikt zijn voor deelname of zal niet willen deelnemen. Dit kan leiden tot fouten.

1.2 Doelstelling

Het voorgaande leidt tot de volgende doelstelling van dit onderzoek.

Het in kaart brengen van de effecten en voor zover mogelijk de omvang van de effecten van non-respons in het Informatienet en het opzetten van een organisatorische structuur beter om te gaan met deze non-respons.

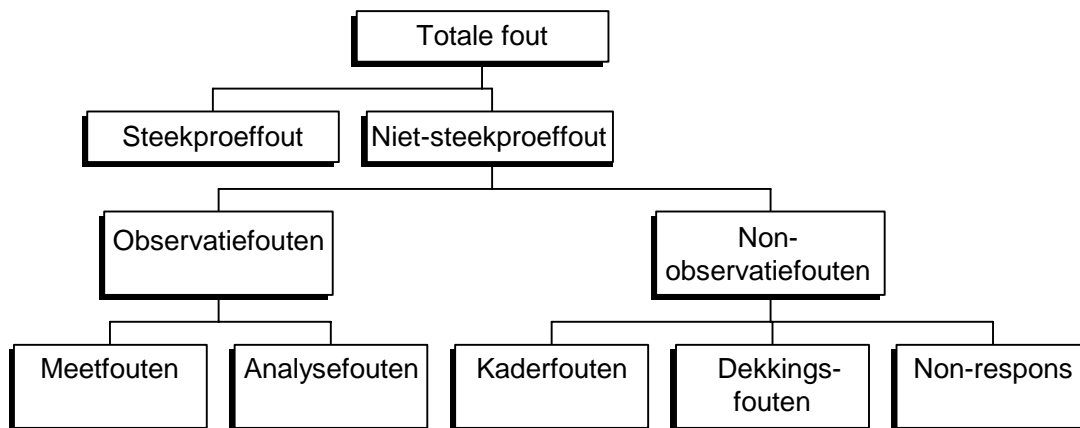
Deelvragen:

1. Wat zijn de consequenties van non-respons op de kwaliteit van de steekproef en op het onderzoek dat op basis van het Informatienet wordt uitgevoerd?
2. Wat zijn de redenen voor non-respons?
3. Welke maatregelen zijn mogelijk om de non-respons te verlagen?

Het optreden van non-respons is niet de enige factor die kan leiden tot fouten. Om het probleem van de non-respons in het juiste kader te beschouwen zal in de volgende paragraaf eerst worden beschreven welke andere factoren kunnen leiden tot fouten. Na deze plaatsbepaling zal in de daaropvolgende paragrafen nader worden ingegaan op de problematiek van de non-respons.

1.3 Fouten in steekproefonderzoek

Het werken met steekproeven kan fouten met zich meebrengen. De verschillende typen van fouten zijn in figuur 1.1 weergegeven.



Figuur 1.1 Typen van fouten in steekproefonderzoek (Zwart, 1994)

Een eerste onderverdeling van de totale fout betreft die in de steekproeffout en de niet steekproeffout. De steekproeffout is een gevolg van het feit dat men op basis van een deel van de populatie uitspraken doet over de gehele populatie (zie paragraaf 1.5 voor een verdere analyse). Afhankelijk van de toevallige selectie van elementen in de steekproef kan de uitkomst van de steekproef anders zijn. Door het hanteren van juiste steekproefprocedures kan de steekproeffout worden gekwantificeerd in de vorm van uitspraken over de betrouwbaarheid van de uitkomsten. Doordat de steekproeffout afhankelijk is van toevalligheden (het toeval bepaalt welke elementen in de steekproef terechtkomen) wordt deze fout ook wel de toevallige, statistische of niet-systematische fout genoemd.

De niet-steekproeffouten hebben een systematische oorzaak en kunnen dus een bias in de onderzoeksresultaten veroorzaken. De niet-steekproeffouten vallen uiteen in de observatie- en de niet-observatiefouten. De observatiefouten bestaan uit meetfouten en analysefouten. Meetfouten kunnen bijvoorbeeld ontstaan door een spraakverwarring tussen de interviewer en de geïnterviewde, maar kunnen ook een gevolg zijn van het bewust geven van verkeerde informatie. Analysefouten wijzen op fouten in de verdere verwerking van de gegevens, bijvoorbeeld in het coderen of interpreteren van gegevens.

De niet-observatiefouten bestaan uit kaderfouten, dekkingsfouten en non-responsfouten. Het steekproefkader voor het Informatienet wordt gevormd door de landbouwtelling. In deze landbouwtelling kunnen fouten zitten. Dekkingsfouten treden op indien er niet een volledige overeenstemming is tussen de populatie waarin men geïnteresseerd is en het steekproefkader op basis waarvan men de steekproef samenstelt. Bedrijven die te laat zijn geteld zullen bijvoorbeeld ontbreken in de landbouwtelling, terwijl deze wel binnen de onderzoekspopulatie vallen. De laatste categorie van fouten zijn de non-responsfouten. Deze categorie zal in de volgende paragraaf nader worden toegelicht.

1.4 Non-responsfouten

Bij het vaststellen van de steekproef wordt de populatie ingedeeld op basis van een aantal

kenmerken en uit deze groepen worden bedrijven getrokken en aangezocht om deel te nemen in het Informatienet. Een deel van deze bedrijven zal echter weigeren deel te nemen. Dit hoeft op zich geen probleem te zijn als deze bedrijven niet afwijken van andere bedrijven. Als de bedrijven die weigeren systematisch bepaalde kenmerken hebben zullen deze kenmerken ondervertegenwoordigd zijn in de steekproef. Als bijvoorbeeld oude agrariërs minder wensen mee te werken leidt dit ertoe dat de leeftijdsverdeling van de deelnemers in het Informatienet niet representatief is voor de leeftijd in de populatie. In hoeverre dit verschil van invloed is op de uitkomsten van een specifiek onderzoek hangt af het van het onderzoeksdoel en in hoeverre de daarvoor relevante kenmerken samenhangen met de leeftijd. De financiering van een bedrijf is bijvoorbeeld bij een oude agrariër heel anders dan bij een jonge beginnende agrariër.

1.5 Steekproeffouten

De steekproeffout wordt veroorzaakt door het feit dat men op basis van een steekproef uitspraken wil doen over de gehele populatie. De schatting op basis van de steekproef is dan afhankelijk van de elementen die (toevallig) in de steekproef worden opgenomen. Indien toevallig een aantal eenheden (bedrijven) met hoge waarden voor de doelvariabelen (bijvoorbeeld inkomen) zijn opgenomen, zal een hogere schatting van het inkomen in de populatie resulteren.

Steekproeffouten hebben als voordeel dat deze gekwantificeerd kunnen worden. Uitgaande van een (gestratificeerde) aselechte steekproef kan een schatting worden gemaakt van de steekproeffout. De steekproeffout oftewel de betrouwbaarheid van de schatting wordt berekend als de wortel uit de variantie van de schatter. De variantie is gelijk aan:

$$V_{\text{ran}} = \left(1 - \frac{n}{N}\right) \frac{S_y^2}{n}$$

waarbij:

$$S_y^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}$$

In geval van een gestratificeerde steekproef, wordt de variantie berekend middels:

$$V_{\text{feit}} = \sum_{h=1}^H W_h^2 \text{var}(\bar{y}_h) = \sum_{h=1}^H W_h^2 (1 - f_h) \frac{S_{yh}^2}{n_h}$$

waarbij:

$$W_h = \frac{N_h}{N} \quad \text{en} \quad f_h = \frac{n_h}{N_h}$$

In beide formules is duidelijk te zien dat de steekproeffout kan worden verkleind door het aantal waarnemingen (n) te vergroten. Des te groter het aantal waarnemingen des te kleiner de kans dat er alleen bedrijven zijn getrokken met een relatief hoog inkomen. Bij twee steekproefbedrijven is dat heel goed mogelijk, bij 20 steekproefbedrijven is het waarschijnlijk dat er zowel bedrijven met hoge als met lage inkomens zijn opgenomen in de steekproef.

1.6 Opzet van onderzoek

In dit rapport wordt met name gekeken naar de non-respons en de effecten daarvan op de representativiteit van de deelnemende bedrijven en op systematische vertekeningen in de resultaten.

In hoofdstuk 2 vindt een analyse plaats van de non-respons. Ten eerste wordt gekeken welke redenen boeren geven om niet deel te nemen. Vervolgens is een analyse gemaakt van de responspercentages in de verschillende groepen. Het vergelijken van responspercentages geeft inzicht in de bereidheid tot meewerken in de verschillende groepen. Hiermee is nog geen inzicht verkregen in het al dan niet optreden van een non-responsbias, laat staan in de omvang van deze bias. Zolang het tweede, derde of zelfs het tiende bedrijf dat uiteindelijk wel mee wil werken niet afwijkt van de bedrijven die daarvoor zijn benaderd is er geen sprake van een selectieve non-respons.

Om dit probleem te ondervangen is vervolgens een kwantitatieve analyse gemaakt om deelnemers te vergelijken met niet-deelnemers. Om te kunnen constateren of de bedrijven die weigeren anders zijn dan bedrijven die meewerken is een koppeling gemaakt met de landbouwtelling. De gegevens uit de landbouwtelling zijn gebruikt om te kijken of er significante verschillen bestaan tussen de bedrijven die weigeren deel te nemen en de bedrijven die zijn geworven. Deze analyse toetst vooral of de structuur van de bedrijven verschilt. In een vervolganalyse met behulp van data-imputatie is gekeken in hoeverre dit ook daadwerkelijk leidt tot een bias in de te schatten doelvariabelen.

In het wervingsproces is geprobeerd mogelijke redenen te verzamelen bij bedrijven die weigeren deel te nemen. Het betreft dan vooral vragen over de mate waarin men op de hoogte is van het LEI en het Informatienet en de attitude ten opzichte van het LEI en de overheid. Dezelfde informatie is ook verzameld bij de bedrijven die wel willen participeren. Ook op basis van deze informatie is vastgesteld of de bedrijven die meewerken significant afwijken van de bedrijven die niet mee willen werken.

Tenslotte is een analyse gemaakt van de representativiteit van het Informatienet door de karakteristieken van de Informatienet-deelnemers te vergelijken met alle bedrijven volgens de landbouwtelling (hoofdstuk 3). Deze analyse is aanvullend doordat het Informatienet een roterend panel is. Ook al is in de aanvullende werving in een specifiek jaar geen sprake van een selectieve non-respons, dan betekent dat niet dat er in voorgaande jaren geen verschillen zijn ontstaan.

2. Analyse non-respons

2.1 Kwalitatieve analyse redenen non-respons

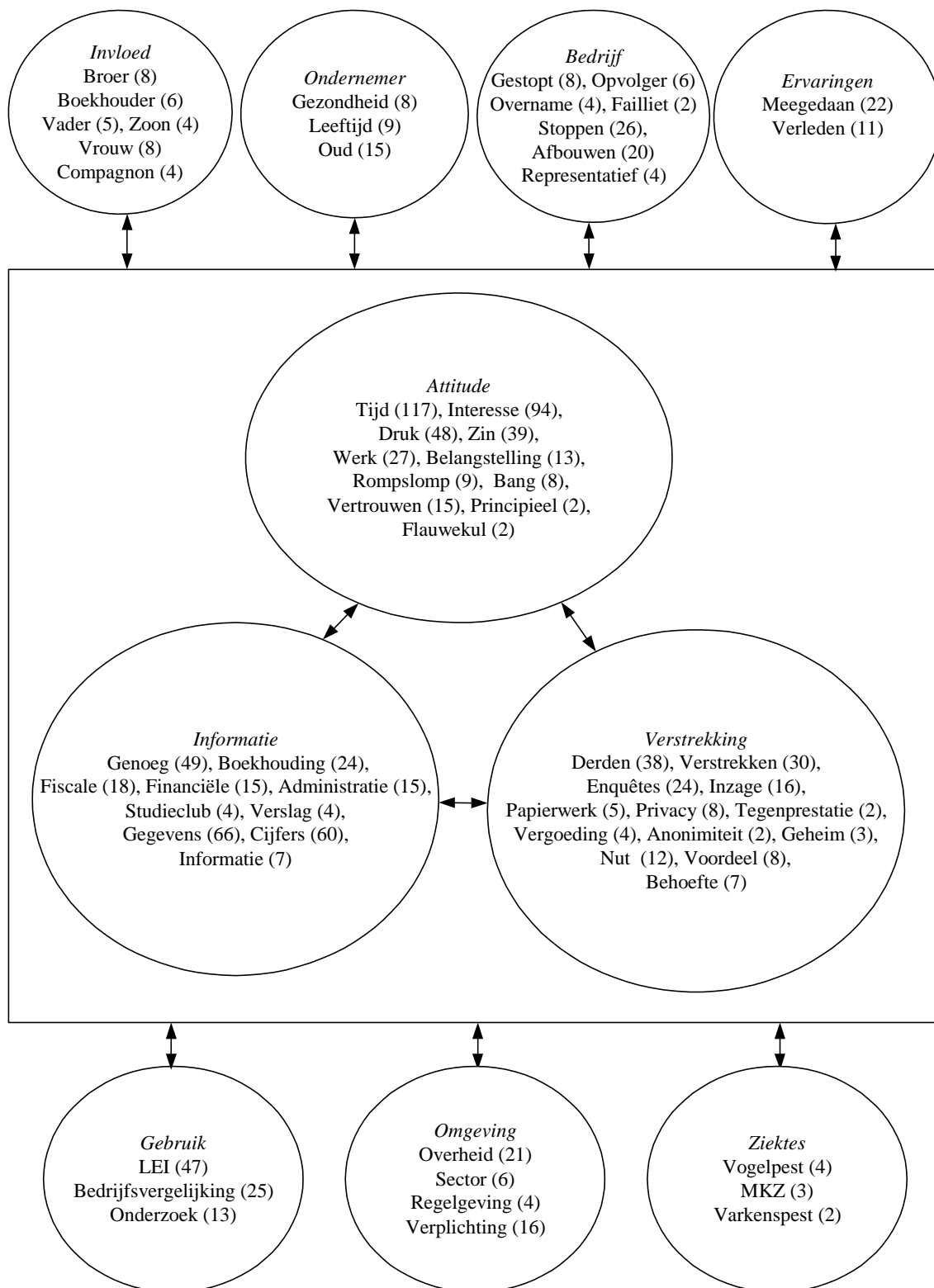
Middels contentanalyse is de inhoud van de opgegeven redenen voor non-respons geanalyseerd. Contentanalyse is een systematische analyse van tekstbestanden (zie bijvoorbeeld: Anders Ericsson en Simon, 1980; Kassarian, 1977; Morris, 1994). In de eerste stap is een telling gemaakt van de begrippen die in de redenen voorkomen. Vervolgens zijn deze begrippen op een logische manier ingedeeld in groepen. De genoemde begrippen, de frequentie en de aangebrachte samenhang is weergegeven in figuur 2.1.

In het centrale blok zijn alle begrippen met betrekking tot de gegevens, de verstrekking van gegevens en de houding van de ondernemer weergegeven. Deze elementen hangen onderling sterk samen. In de cirkel 'attitude' zijn een aantal begrippen opgenomen die vaak worden genoemd als reden voor non-respons. Gebrek aan tijd, te veel werk, geen interesse en geen belangstelling worden door veel boeren als reden voor weigering naar voren gebracht. Daarnaast zijn er een aantal termen die duiden op het vertrouwen in een zorgvuldige behandeling van hun gegevens.

In de cirkel 'verstrekking' zijn een aantal begrippen terug te vinden die duiden op argumenten voor het niet verstrekken van gegevens. Een substantieel aantal boeren heeft er moeite mee gegevens te verstrekken of inzage te geven aan derden. Privacy en anonimiteit zijn daarbij belangrijke termen. Een aantal boeren is wat pragmatischer ingesteld en benadrukt de kosten-batenverhouding. Het levert hen extra papierwerk op en er staat geen enkele of onvoldoende tegenprestatie tegen over. Het nut van deelname en daarmee de behoefte aan deelname wordt door deze boeren als beperkt ervaren. Een aantal boeren zegt bereid te zijn gegevens te verstrekken als er voor betaald zou worden.

In de cirkel 'informatie' staan een aantal begrippen gerelateerd aan de huidige informatie. Deze begrippen worden vaak genoemd in samenhang met de overige twee cirkels. Boeren zeggen al voldoende cijfers te hebben uit hun boekhouding, administratie of studieclub. Het nut van extra informatie wordt niet herkend. Daarnaast heeft men moeite inzage te geven in de gegevens uit de boekhouding.

Naast deze 3 kerncirkels zijn een aantal groepen begrippen relevant voor het niet deelnemen aan het informatienet. Ten eerste wordt vaak als reden voor niet deelnemen een andere persoon aangevoerd (samengevat in de groep 'invloed'). De broer, vader of vrouw is het vaak niet met deelname eens of zou het niet met deelname eens zijn (het betreft hier de opgave van de boer die het zo percipieert of zich daarachter verschuilt). Ook wordt een aantal keren het negatieve advies van de boekhouder aangevoerd. Daarnaast is er een groep begrippen die aan de situatie van de ondernemer refereren. De ondernemer brengt zijn gezondheid of leeftijd naar voren als reden voor niet deelname. Daarnaast hebben een aantal boeren ervaringen uit het verleden. Voor sommigen heeft dit geleid tot negatieve ervaringen, voor anderen is dit een reden om de administratieve last nu maar bij iemand anders neer te leggen.



Figuur 2.1 Analyse van begrippen die een rol spelen bij de non-respons

Een groot aantal boeren draagt argumenten aan op basis van de toestand van het bedrijf (groep 'bedrijf'). Het betreft met name bedrijven die zijn gestopt of binnenkort gaan stoppen of afbouwen. Een aantal boeren vindt zijn eigen bedrijf niet representatief. Dit loopt uiteen van een boer die zijn eigen bedrijf te goed vindt en daarmee bang is de resultaten (te) positief te beïnvloeden, tot boeren die vinden dat ze een uniek bedrijf hebben dat toch nergens mee is te vergelijken.

De begrippen onder het kader vinden hun oorsprong buiten het eigen bedrijf. De eerste groep van begrippen met een oorsprong buiten het eigen bedrijf, maar een duidelijke impact op het eigen bedrijf, zijn de besmettelijke dierziektes die worden genoemd. MKZ, varkenspest en vogelpest werden genoemd als redenen. Van de omgevingsfactoren wordt de overheid vaak genoemd. Gebrek aan vertrouwen in de overheid wordt als reden genoemd om niet deel te nemen. Een aantal boeren beschuldigt de overheid van 'het de nek om draaien van de agrarische sector'. Overmatige regelgeving van de overheid wordt ook een aantal keren expliciet naar voren gebracht. Als reactie op al deze regelgeving en verplichtingen wil men niet aan het Informatienet meedoen indien dit niet verplicht is. Men is moe van allerlei enquêtes en vragenlijsten. De weerzin tegen de overheid keert zich in die gevallen ook tegen het LEI. Veel boeren maken geen onderscheid tussen het LEI en de overheid. Het LEI wordt door sommige boeren gezien als een verlengstuk van de overheid. Deze boeren geven aan dat het gebruik van de gegevens niet ten goede komt van de agrarische sector.

2.2 Kwantitatieve analyse non-respons

Tabel 2.1 geeft de responspercentages in de verschillende bedrijfstypen. De verschillen in respons lopen duidelijk uiteen. In bijvoorbeeld de akkerbouw is meer dan 50% van de geschikte bedrijven bereid deel te nemen. Ook in de melkveehouderij is de bereidwilligheid relatief hoog. Dit zijn belangrijke sectoren in de Nederlandse landbouw en voor het beleid. Daar tegenover zijn er ook sectoren die erg laag scoren. Van de vleesvarkenbedrijven en gesloten varkensbedrijven wil maar 20% van de geschikte bedrijven deelnemen. In de tuinbouw is een zeer lage respons waar te nemen voor de snijbloemenbedrijven (17,4%). Het gemiddelde van het totaal aantal benaderde bedrijven ligt op 32%. Dit percentage wordt natuurlijk sterk naar beneden gedrukt door de slecht responderende groepen. Een gewogen gemiddelde naar het belang van de verschillende sectoren in de Nederlandse land- en tuinbouw zou hoger uitvallen.

In tabel 2.2 is een opsplitsing gegeven naar de responspercentages in de verschillende grootteklassen van het bedrijf (zie Vrolijk en Van der Veen, 2004, voor een beschrijving van de klassen). Op basis van de vorige tabel met de lage respons in vooral de tuinbouw zou men de verwachting kunnen hebben dat de respons in de hoogste klasse lager zou zijn. Uit de cijfers blijkt dit echter nauwelijks. De respons in de kleinste en grootste klasse is net iets lager dan in de middelste klasse. Het percentage ongeschikte bedrijven is duidelijk groter in de kleinste grootteklasse. Dit hangt samen met het feit dat bedrijven die aan het afbouwen zijn, en hun bedrijf op korte termijn van de hand doen, vaak in de kleinste klasse vallen.

Tabel 2.1 Resultaten aanvullende werving deelnemers informatienet voorjaar 2003 naar bedrijfstype

Bedrijven	Be- naderd	Ge- weigerd	Ge- worven	On- geschikt	Geworven % van bena- derd	Geworven % van ge- schikt	% On- geschikt
Akkerbouw	27	11	12	4	44,4	52,2	14,8
Fabrieksaardappelen	35	15	10	10	28,6	40,0	28,6
Akkerbouw biologisch	41	17	13	11	31,7	43,3	26,8
Melkvee	95	47	33	15	34,7	41,3	15,8
Melkvee biologisch	5	1	4		80,0	80,0	0,0
Klavermeisterijen	67	31	17	19	25,4	35,4	28,4
Andere graasdieren	82	37	14	31	17,1	27,5	37,8
Fokvarkens	57	28	14	15	24,6	33,3	26,3
Vleesvarkens	145	94	26	25	17,9	21,7	17,2
Gesloten varkens	56	39	12	5	21,4	23,5	8,9
Legkippen	11	5	5	1	45,5	50,0	9,1
Vleespluimvee	10	4	4	2	40,0	50,0	20,0
Ander hokdieren	40	21	16	3	40,0	43,2	7,5
Combinaties	126	77	30	19	23,8	28,0	15,1
Gewascombinaties biologisch	26	11	9	6	34,6	45,0	23,1
Glasgroente	111	45	38	28	34,2	45,8	25,2
Snijbloemen	82	38	8	36	9,8	17,4	43,9
Planten	11	6	2	3	18,2	25,0	27,3
Overige glas	70	33	15	22	21,4	31,3	31,4
Opengrondsgroente	70	44	14	12	20,0	24,1	17,1
Fruit	137	68	28	41	20,4	29,2	29,9
Boomkwekerij	61	29	16	16	26,2	35,6	26,2
Paddestoel	33	16	8	9	24,2	33,3	27,3
Bloembollen	76	47	16	13	21,1	25,4	17,1
Overige opengrond	80	44	14	22	17,5	24,1	27,5
TOTAAL	1.554	808	378	368	24,3	31,9	23,7

Tabel 2.2 Resultaten aanvullende werving 2001 naar grootteklasse van het bedrijf

Bedrijven	Be- naderd	Ge- weigerd	Ge- worven	On- geschikt	Geworven % van bena- derd	Geworven % van ge- schikt	% On- geschikt
Kleinste ege-klasse	618	285	131	202	21,2	31,5	32,7
Middelste ege-klasse	277	151	78	48	28,2	34,1	17,3
Grootste ege-klasse	659	372	169	118	25,6	31,2	17,9

In tabel 2.3 zijn de resultaten weergegeven van de stellingen die aan de benaderde boeren zijn voorgelegd. Uit deze tabel blijkt een duidelijk verschil tussen de deelnemers en de niet-deelnemers. Zo zijn de deelnemers beter op de hoogte van het LEI en het Informatienet. Ook geldt dat de deelnemers beter op de hoogte zijn van het gebruik van het Informatienet. De niet-deelnemers zijn het gemiddeld oneens met de stelling of ze op de hoogte zijn van het

gebruik van het Informatienet. Het Informatienet en het leveren van informatie voor het Informatienet wordt door de deelnemers als nuttiger ervaren. Ten aanzien van de zorgvuldigheid en de objectiviteit van het LEI hebben de deelnemers een positievere indruk dan de niet-deelnemers. De laatste stelling laat een duidelijk verschil zien. De niet-deelnemers hebben een significant lager vertrouwen in de overheid.

Tabel 2.3 Houding agrariërs naar deelnemers en niet-deelnemers (-2 oneens tot 2 eens)

	Niet deelnemers		Deelnemers	
	gemiddelde	SF	gemiddelde	SF
Op de hoogte van het LEI	0,9	0,05	1,3	0,07
Op de hoogte van het Informatienet	0,0	0,05	0,9	0,07
Op de hoogte gebruik Informatienet	-0,4	0,05	0,5	0,07
Nut Informatienet	0,1	0,04	0,8	0,06
Nut verstrekken gegevens	0,1	0,04	0,9	0,06
Zorgvuldigheid LEI	0,2	0,03	0,7	0,06
Objectiviteit LEI	0,2	0,03	0,6	0,06
Vertrouwen overheid	-0,7	0,04	0,3	0,06

SF - standaardfout

Indien het al dan niet willen deelnemen wordt verklaard uit de score op de hiervoor weergegeven stellingen blijkt dat de stelling ten aanzien van het vertrouwen in de overheid het grootste voorspellende vermogen heeft. De stelling die als tweede het grootste voorspellende vermogen heeft is de mate waarin men op de hoogte is van het gebruik van het Informatienet. Het gepercipieerde nut van het leveren van gegevens aan het Informatienet is de derde verklarende factor die van invloed is op het al dan niet deelnemen.

De eerste factor 'vertrouwen in de overheid' is moeilijk door het LEI te beïnvloeden. Het LEI beschikt niet over sturingsmogelijkheden om dit te beïnvloeden. Wel kan dit aanleiding zijn om de onafhankelijke positie van het LEI meer naar voren te brengen. Voor de tweede factor 'op de hoogte van het gebruik van het Informatienet' heeft het LEI deze sturingsmogelijkheden wel. Middels een betere communicatie via bijvoorbeeld vakbladen kan ervoor gezorgd worden dat potentiële Informatienet deelnemers beter op de hoogte worden gesteld van de manier waarop het Informatienet wordt gebruikt.

2.3 Kwantitatieve analyse non-responsbias

Op basis van een analyse van de kenmerken van bedrijven kan gekeken worden in hoeverre responderende bedrijven afwijken van niet-responderende bedrijven. Dit geeft een inzicht in het optreden van een non-respons maar niet in het effect op de te schatten variabelen.

Middels data-imputatie (zie Vrolijk, 2004 and Vrolijk and Dol, 2003) is een inschatting gemaakt van het effect van de non-respons op de te schatten doelvariabelen. Hierbij is vooral

gekeken naar de melkveehouderij omdat de aantallen waarnemingen daar dusdanig groot zijn dat er zinvolle uitspraken te doen zijn.

Op basis van het aantal melkkoeien, het aantal hectares cultuurgrond, het aantal hectares voedergewassen en de totale omvang van het bedrijf is een inschatting gemaakt van een set van doelvariabelen. Deze inschatting is gemaakt voor zowel de bedrijven die deelnemen als de bedrijven die niet deelnemen. Voor de schatting is gebruik gemaakt van de bedrijven die voor de werving al aan het Informatienet deelnamen en de bedrijven die in het betreffende jaar voor het eerst zijn uitgewerkt. De resultaten zijn weergegeven in tabel 2.4.

Tabel 2.4 Non-responsbias voor de melkveehouderij (kleinste ege-klasse)

	Geweigerd		Geworven		Ongeschikt	
	gemiddelde	Std. fout	gemiddelde	Std. fout	gemiddelde	Std. fout
<i>Imputatie-variabelen:</i>						
Koeien	34,3	1,7	38,0	2,1	25,0	3,2
Ha cultuur	22,7	1,3	26,1	1,7	19,9	2,3
Ha voeder	22,1	1,1	25,2	1,5	19,0	2,2
EGE	65,0	2,9	72,8	3,5	50,6	5,2
<i>Doelvariabelen:</i>						
Gezinsinkomen uit bedrijf	28.486	4.085	31.815	4.048	19.223	5.569
Besparingen	5.735	3.947	10.737	3.957	-1.889	4.102
Melkoprangsten	82.894	6.201	102.573	8.727	60.232	9.356
N mutatie	360,5	161,1	666,2	323,7	410,2	195,4
Bedrijfsresultaat	-37.889	3.734	-36.159	4.815	-40.689	5.346
Opbrengst per €100 kosten	71,5	3,2	76,6	3,4	66,8	2,7
Opbrengst per AJE	13.647	3.415	17.194	3.454	4.728	3.758

Tabel 2.4 geeft een indicatie van de omvang van de bias. De eerste regels in de tabel geven het verschil in de structuur van de bedrijven. Uit deze gegevens blijkt dat de geworven bedrijven voor alle weergegeven kenmerken iets groter zijn dan de bedrijven die weigeren. In alle gevallen zijn de verschillen niet significant. Opmerkelijk is het verschil tussen de bedrijven die als ongeschikt worden aangemerkt en de bedrijven die wel of niet meedoen. De bedrijven die als ongeschikt worden aangemerkt zijn voor een aantal kenmerken significant kleiner dan de overige bedrijven.

Interessanter is om te zien in hoeverre deze structuur kenmerken zich vertalen in een bias in de te schatten doelvariabelen. In de tweede helft van de tabel zijn de schattingen weergegeven. De verschillen in structuur tussen de deelnemende en niet-deelnemende bedrijven vertalen zich in lichte verschillen in de resultaten. Deze verschillen in resultaten zijn echter niet significant. Wederom geldt dat de verschillen met de ongeschikte bedrijven groter zijn dan tussen de wel en niet-deelnemende bedrijven. Voor een aantal kengetallen bestaat er een significant verschil tussen de deelnemers en de ongeschikte bedrijven. Op basis van deze resultaten kan voorzichtig de conclusie worden getrokken dat het al dan niet als ongeschikt bestempelen van een bedrijf meer invloed op de bias heeft dan voorheen gedacht. Het effect

van de non-respons lijkt te bestaan, maar is niet significant. Deze inschatting van het mogelijke effect is uitsluitend gebaseerd op de waar te nemen variabelen. Hiermee is nog geen uitspraak te doen over niet waargenomen variabelen, bijvoorbeeld de kwaliteit van het management. Indien de deelnemende bedrijven een veel hoger (of juist lager) kwaliteitsniveau van management hebben, kan dit nog steeds leiden tot een significant effect. Deze variabele zal eerst nader moeten worden geoperationaliseerd en worden gemeten voordat daar zinvolle uitspraken over gedaan kunnen worden.

2.4 Ongeschikt zijn van bedrijven voor deelname aan Informatienet

Een deel van de bedrijven worden door de TAM als ongeschikt aangemerkt. De eerste reden hiervoor is het onbereikbaar zijn van de boer. Daarnaast lijkt ook de meest recente landbouwtelling soms al verouderd te zijn.

De overblijvende begrippen zijn wederom geanalyseerd middels contentanalyse. De begrippen zijn geteld en ingedeeld in groepen. In figuur 2.2 is de belangrijkste groep van begrippen in het centrum weergegeven. Al deze begrippen verwijzen naar een situatie waarin het bedrijf is beëindigd of aan het afbouwen is. Deze bedrijven kunnen terecht als niet geschikt worden bestempeld. Het is natuurlijk wel de vraag hoeveel bedrijven in de landbouwtelling dan ook niet meer bestaan. Middels de bedrijven die zijn benaderd voor deelname zou daarvan een schatting gemaakt kunnen worden.

De groep 'bedrijf' geeft aan dat bepaalde benaderde bedrijven in de ogen van de ondernemer of TAM geen agrarisch bedrijf zijn. Er zijn bedrijven benaderd die feitelijk een accountantsbureau of inrichting zijn. Ook dit roept vragen op omtrent de omvang van dit probleem in de landbouwtelling. De groep 'neven' duidt op situaties waarin wel sprake is van agrarische activiteiten maar waarbij dit slechts een deel van de totale economische activiteit is of dat het niet echt als economische activiteit wordt verricht maar meer als een hobby. Het weglaten van deze groep van bedrijven is mede de oorzaak van het in de vorige paragraaf beschreven verschijnsel dat de kleine bedrijven ondervertegenwoordigd zijn. De grote financiële verwevenheid zou een praktische reden kunnen zijn om dergelijke bedrijven buiten het Informatienet te houden. Deze reden komt duidelijk naar voren in de groep 'complexiteit'. Verschillende locaties, verwevenheid met activiteiten in en met het buitenland kunnen ertoe leiden dat de kosten en opbrengsten niet eenduidig aan de agrarische activiteit zijn toe te wijzen.

Daarnaast zijn er nog een aantal kenmerken gerelateerd aan de ondernemer of het bedrijf die ertoe leiden dat een bedrijf als ongeschikt wordt aangemerkt. De kenmerken van de ondernemer zijn bijvoorbeeld de leeftijd of de gezondheid. De begrippen van het bedrijf zijn verhuurd, fiscaal en representatief. Verhuurd impliceert dat er nog wel sprake is van agrarische activiteit maar dan door een andere partij. Het is dan de vraag in hoeverre die partij benaderd zou moeten worden. Fiscaal refereert aan situaties waarin de boer al aangeeft dat er een bedrijfsopzet is gekozen vooral met het doel om er fiscaal gunstig van af te komen. Dit kan het in administratie nemen van een dergelijk bedrijf bemoeilijken. Het niet-representatief zijn van een bedrijf is wetenschappelijk een rare reden en duidt op onvoldoende kennis bij potentiële deelnemers en TAM's op dit vlak. Een bedrijf op zich kan nooit representatief zijn. Een groep van bedrijven en juist de verscheidenheid binnen deze groep kan representatief zijn

voor de populatie.

De laatste groep van begrippen is genaamd 'activiteiten'. Ten dele beschrijven deze begrippen de activiteiten die voorkomen op de nevenbedrijven en ten dele worden de begrippen opgevoerd als reden voor ongeschikt zijn. Dit impliceert dat de TAM een beoordeling maakt van de representativiteit van een geselecteerd bedrijf. Door deze beoordeling wordt eigenlijk een deel van de bedrijven die in de landbouwtelling voorkomen uitgesloten van een mogelijke deelname aan het Informatienet.

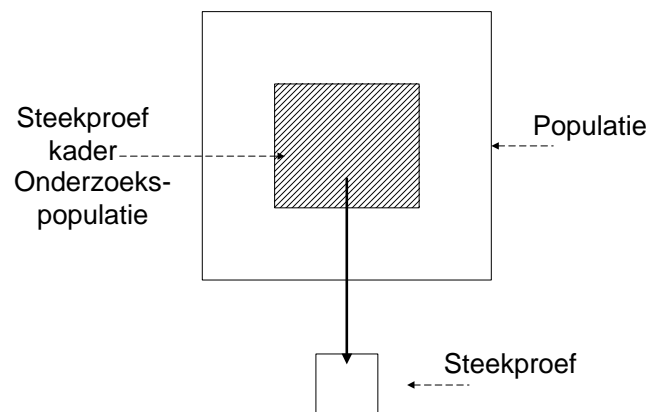


Figuur 2.2 Analyse van begrippen die een rol spelen bij het ongeschikt zijn van het bedrijf

3. Representativiteit Informatienet

3.1 Definitie van representativiteit

Voor de beoordeling van de representativiteit is het onderscheid tussen populatie en onderzoekspopulatie van groot belang (zie figuur 3.1). De populatie wordt bijvoorbeeld gevormd door alle agrarische bedrijven in Nederland. Voor het Informatienet zijn de grenzen van 16 en 1.200 ege gesteld. De onderzoekspopulatie wordt dus gevormd door alle bedrijven in Nederland tussen 16 en 1.200 ege. Deze bedrijven zijn opgenomen in de landbouwtelling. De bedrijven in de landbouwtelling tussen 16 en 1.200 ege vormen het steekproefkader. Op basis waarvan de steekproef wordt getrokken. Door de bijna volmaakte overeenstemming tussen de onderzoekspopulatie en het steekproefkader (er ontbreken wellicht een paar bedrijven die te laat zijn geteld) is er nauwelijks sprake van dekkingfouten.



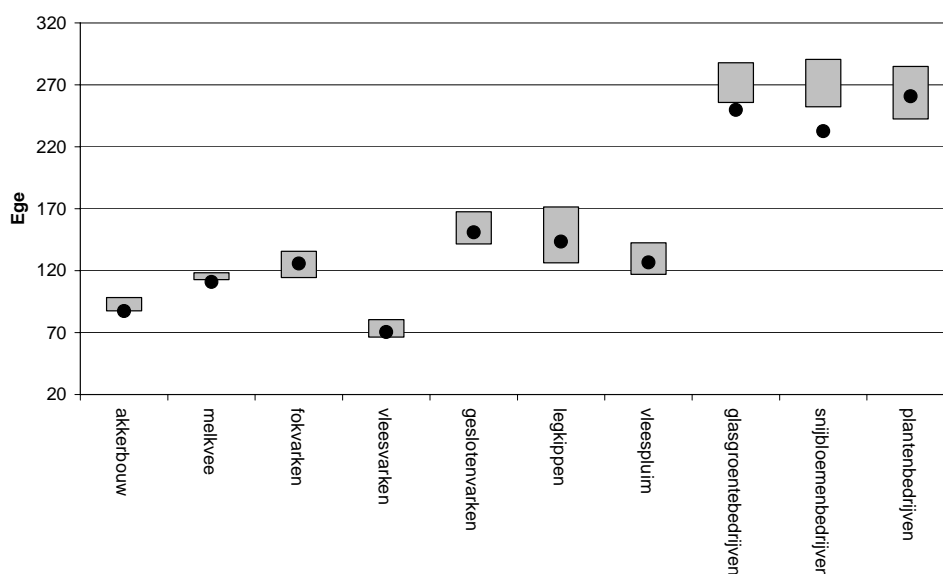
Figuur 3.1 Populatie en steekproef

Voor het beoordelen van de representativiteit is het niet zinvol om de onderzoekspopulatie met de gehele populatie te vergelijken. Uitspraken zoals 'het Informatienet is representatief voor 75% van de bedrijven omdat 75% van de bedrijven in de range van 16 en 1.200 ege vallen' zijn dus niet gegrond. Een dergelijke uitspraak dient vervangen te worden door bijvoorbeeld 'het Informatienet geeft een beeld van of representeert 75% van de bedrijven uit de landbouwtelling' (zie Vrolijk et al., 2004, tabel 5.2 voor een vergelijking van de steekproef en de populatie). Voor het beoordelen van de representativiteit van de steekproef, moet de steekproef worden vergeleken met de onderzoekspopulatie, bijvoorbeeld door te kijken of het percentage akkerbouwbedrijven in de onderzoekspopulatie even groot is als het percentage akkerbouwbedrijven in de steekproef.

3.2 Representativiteit van het Informatienet

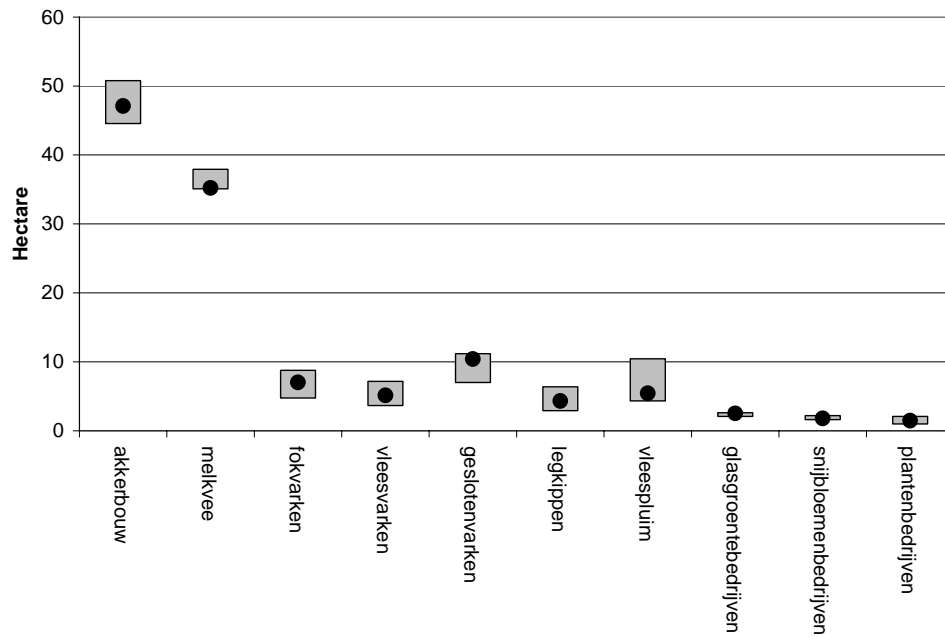
In diverse publicaties wordt gesteld dat het Informatienet representatief is voor de onderzoekspopulatie. Een dergelijke algemene opmerking is moeilijk hard te maken. Een steekproef kan representatief zijn ten aanzien van een variabele maar niet representatief ten aanzien van een andere. Een steekproef zal nooit representatief zijn ten aanzien van alle variabelen, tenzij de steekproef bijna samenvalt met de gehele populatie. Het vraagstuk van de representativiteit van de steekproef moet dus voor elk onderzoek afzonderlijk worden beoordeeld. De representativiteit moet dan worden beoordeeld ten aanzien van de variabelen die in dat onderzoek relevant zijn. Indien de leeftijd van de ondernemer geen invloed heeft op de te onderzoeken variabele is het niet zo belangrijk dat de steekproef representatief is ten aanzien van dit aspect. In een onderzoek waarin dit wel belangrijk is (bijvoorbeeld bedrijfsopvolging) dient de representativiteit ten aanzien van leeftijd te worden getoetst.

In figuur 3.2 is deze vergelijking gemaakt voor de bedrijfsomvang uitgedrukt in ege. De grijze blokjes geven de grenzen van het 95% betrouwbaarheidsinterval voor de schatting op basis van het Informatienet weer. De schatting is gemaakt voor elk afzonderlijk type. De stip geeft het daadwerkelijke gemiddelde volgens de landbouwtelling weer. Als er geen significant verschil bestaat tussen het gemiddelde volgens het Informatienet en het gemiddelde volgens de landbouwtelling moet deze stip in het grijze blokje vallen. Bij inspectie van de figuur blijkt dat de stip zich voor bijna alle types binnen het betrouwbaarheidsinterval valt.¹ Dit betekent dat de schatting op basis van het Informatienet geen significante afwijking van het aantal ege laat zien.

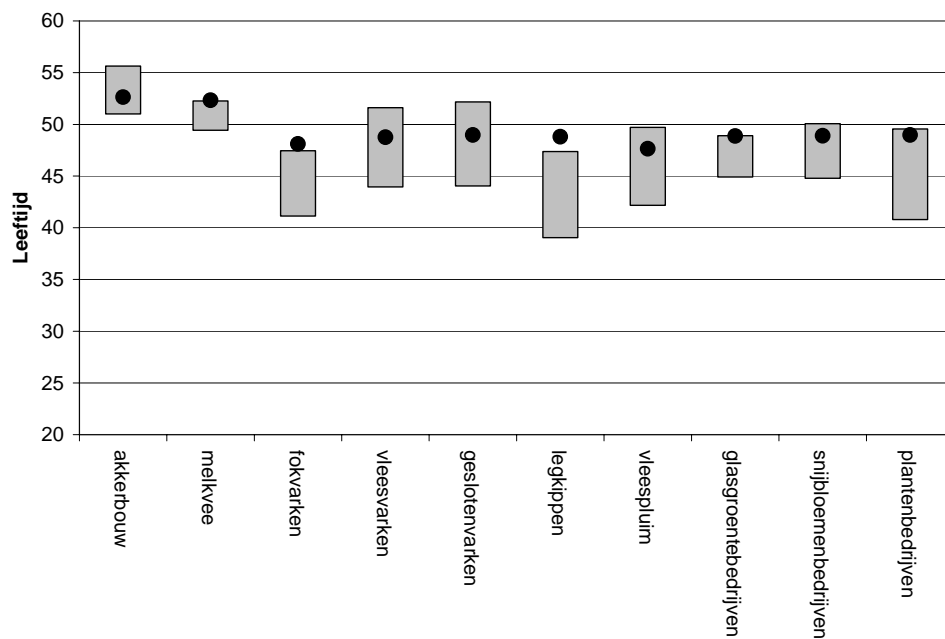


Figuur 3.2 Representativiteit Informatienet voor het aantal ege

¹ Dit geldt niet voor de glasgroentebedrijven, de snijbloemenbedrijven en plantenbedrijven. Dit wordt veroorzaakt doordat dit geen typen zijn zoals in de stratificatie opgenomen.



Figuur 3.3 Representativiteit Informatienet voor de oppervlakte van een bedrijf

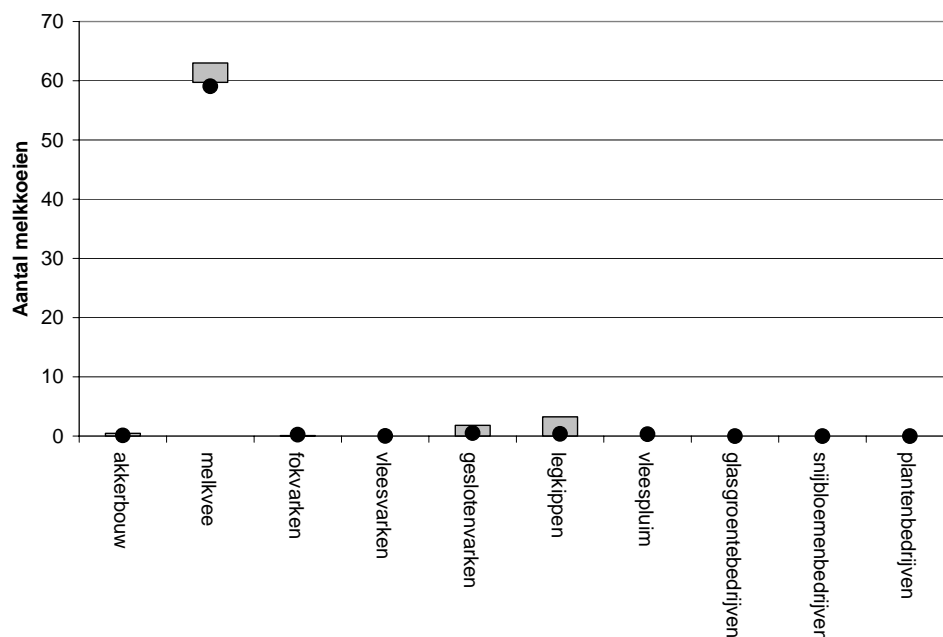


Figuur 3.4 Representativiteit Informatienet voor leeftijd

In figuur 3.3 is een soortgelijke analyse uitgevoerd, maar dan naar de oppervlakte van het bedrijf. De blokjes geven wederom de grenzen van het 95% betrouwbaarheidsinterval van de schatting op basis van het Informatienet weer en de stip geeft het daadwerkelijke gemiddelde volgens de landbouwtelling. Uit deze figuur blijkt dat het Informatienet tot schattingen van het oppervlakte leidt die niet significant afwijken van de landbouwtelling voor het overgrote deel van de bedrijfstypen.

In figuur 3.4 is de representativiteit van het Informatienet voor leeftijd weergegeven. De representativiteit ten aanzien van leeftijd is aanzienlijk lager dan die ten aanzien van oppervlakte en ege-omvang. De schatting op basis van het Informatienet heeft de neiging een te lage leeftijd te hebben. Voor een aantal types zoals fokvarkensbedrijven en legkippenbedrijven resulteert het Informatienet in een significante onderschatting van de leeftijd.

De voorgaande figuren geven informatie over kenmerken die voor alle bedrijfstypen relevant zijn. In specifiek onderzoek kan de representativiteit van belangrijke variabelen worden onderzocht. In de onderstaande figuur is bijvoorbeeld getoetst of het gemiddeld aantal melkkoeien volgens het Informatienet afwijkt van dat volgens de landbouwtelling. Uit de figuur blijkt dat het aantal melkkoeien op melkveebedrijven in het Informatienet iets hoger ligt dan het aantal melkkoeien op melkveebedrijven volgens de landbouwtelling.



Figuur 3.5 Representativiteit Informatienet voor het aantal melkkoeien

Uit deze figuren blijkt dat het Informatienet in grote mate representatief is voor de belangrijke stratificatievariabele ege. Ook voor oppervlakte resulteert het Informatienet in goede schattingen. Voor leeftijd geldt dat voor een aantal typen het Informatienet resulteert in een significante onderschatting.

4. Conclusies en aanbevelingen non-respons

De analyse van de non-respons heeft betrekking op de werving voor het boekjaar 2001. Deze werving is pas in het voorjaar 2003 uitgevoerd. Bij de interpretatie van de uitkomsten is het van belang het specifieke karakter van deze werving in beschouwing te nemen. De werving was vooral gericht op bedrijven voor de EU-variant (zie Vrolijk et al., 2004). Daarnaast geldt dat de werving vooral telefonisch is uitgevoerd. De uitkomst van de werving kan hierdoor zijn beïnvloed.

In de analyse van de non-responsbias is geconstateerd dat er een licht verschil bestaat tussen de responderende en niet-responderende bedrijven. Deze verschillen zijn echter niet significant. Opmerkelijk is wel dat de ongeschikte bedrijven zowel qua structuur als qua resultaten significant afwijken van de andere twee groepen. In samenhang met de constatering dat het percentage ongeschikte bedrijven sterk uiteenloopt tussen verschillende wervers leidt dit tot de aanbeveling de criteria voor het al dan niet geschikt zijn van bedrijven duidelijker te formuleren en te communiceren.

Uit de analyse van de vragen omtrent de mening van de benaderde boeren blijkt dat vooral het vertrouwen in de overheid en de mate waarin de boer op de hoogte is van het gebruik van het Informatienet bepalen of een boer al dan niet deelneemt. De eerste factor 'vertrouwen in de overheid' is moeilijk door het LEI te beïnvloeden. Voor de tweede factor 'op de hoogte van het gebruik van het Informatienet' heeft het LEI deze wel. Middels een betere communicatie via bijvoorbeeld vakbladen kan ervoor gezorgd worden dat potentiële Informatienet deelnemers beter op de hoogte worden gesteld van de manier waarop het Informatienet wordt gebruikt.

Uit de analyses omtrent de representativiteit van het Informatienet blijkt dat het Informatienet in grote mate representatief is voor de belangrijke stratificatievariabele ege. Ook voor oppervlakte resulteert het Informatienet in goede schattingen. Voor variabelen die verder afstaan van de variabelen die bij de opzet van het Informatienet zijn gebruikt, moet in een specifiek onderzoek worden nagegaan of de opgenomen bedrijven inderdaad representatief zijn.

Uit de gegevens blijkt verder dat er grote verschillen bestaan tussen de TAM's die bedrijven werven (vanwege de privacy zijn deze resultaten niet opgenomen in dit rapport). De verschillen zitten zowel in de responspercentages als het percentage bedrijven dat als ongeschikt voor deelname aan het Informatienet wordt aangemerkt. Ten dele kunnen deze verschillen worden verklaard uit de bedrijfstypen waarmee een TAM met name werkt. Er is echter wel degelijk een persoonseffect. Bij toekomstige wervingen verdient het aanbeveling met name die TAM's te laten werven die er lol aan hebben, de betekenis van het Informatienet duidelijk kunnen communiceren en zelf overtuigd zijn van het nut van het Informatienet. De effectiviteit van deze groep van TAM's kan verder worden verhoogd door gericht trainingen te geven. Deze training kan enerzijds betrekking hebben op het wervingsproces en de interactie met de deelnemer, maar anderzijds ook op onduidelijke aspecten zoals redenen voor het ongeschikt zijn en de zogenaamde representativiteit van een bedrijf.

Uit de analyse van het niet geschikt zijn van bedrijven blijkt dat elk jaar een groot deel van de bedrijven als ongeschikt voor deelname aan het Informatienet wordt bestempeld. Ten dele wordt dit veroorzaakt door praktische problemen zoals het herhaaldelijk onbereikbaar zijn van een bedrijf. Daarnaast geldt dat veel bedrijven ondertussen zijn gestopt. Dit maakt de bedrijven inderdaad ongeschikt voor deelname aan het Informatienet, maar roept wel de vraag op hoeveel soortgelijke bedrijven er in de landbouwtelling voorkomen. Er zijn echter ook meer subjectieve aspecten waarom een bedrijf als ongeschikt wordt bestempeld. Nevenactiviteiten of het ondergeschikt zijn van de agrarische activiteiten leiden er vaak toe dat een bedrijf niet wordt opgenomen. Systematisch (kleine) bedrijven als niet ter zake doende af te wijzen, leidt tot een slechte vertegenwoordiging van deze bedrijven. Het is te overwegen deze agrarische activiteiten wel in het Informatienet op te nemen, of de 16 ege grens moet worden herzien. Verder kan de complexiteit deels worden opgelost door deze bedrijven in de EU-variant op te nemen en niet alle data per producteenheid te verzamelen, maar alleen de bedrijftotalen te verzamelen.

In het verleden zijn ook analyses van de non-respons gemaakt, maar dan over tijdreeksen van vier of vijf jaar. De resultaten van de werving zijn afhankelijk van stemmingen op een bepaald moment. Door de veranderingen van de werving en de verbeterde registratie van de non-repons is deze analyse slechts gebaseerd op gegevens van één jaar. Het verdient aanbeveling om na enkele jaren een soortgelijke analyse uit te voeren over de werving van een aantal jaren.

5. Deel II: Rotatiemethodieken - Inleiding en doelstelling

5.1 Inleiding

In het Informatienet wordt jaarlijks een deel van de bedrijven vervangen. Vervanging en aanvulling is noodzakelijk in verband met de uitval sinds de vorige werving, het verversen van de steekproef om structuurveranderingen in de Nederlandse landbouw te volgen (Dijk, 1989) en de geldende maximumtermijn voor participatie van bedrijven in het Informatienet. Deze maximumtermijn is door het LEI zelf ingesteld en hoort bij de huidige rotatiemethode. Daarnaast geldt dat de samenstelling van de primaire agrarische bedrijven in Nederland jaarlijks een ontwikkeling doormaakt. Daarmee veranderen ook de bedrijven in de Landbouwtelling en dus de bedrijven in het steekproefkader.

Voor 1974 was het uitgangspunt dat bedrijven in de regel niet langer dan vijf jaar in administratie werden gehouden, tenminste als er geen sprake was van uitval. In 1974 is dit aantal verhoogd naar 6 à 7 jaar. Omdat jaarlijks ongeveer 10% van de bedrijven die in administratie zijn, wegens uitval niet voor een volgend boekjaar beschikbaar is, werd in de praktijk ieder jaar ongeveer 25% van de bedrijven vervangen door nieuwe bedrijven die steekproefsgewijs getrokken werden (Dijk, 1989). Door de vernieuwing in het boekhoudsysteem is deze termijn momenteel nog verder verruimd.

Het Informatienet heeft de vorm van een roterend panel. Een panel is een combinatie van een cross-sectie en een tijdreeks. Een cross-sectie is een dwarsdoorsnede in de tijd. Dit betekent dat er gekeken wordt naar gegevens over een aantal verschillende variabelen, objecten of entiteiten op één tijdstip of in een (vaste) periode. Een tijdreeks bevat gegevens over één variabele op een aantal tijdstippen. Bij een panel worden gegevens van meerdere variabelen op een aantal verschillende tijdstippen beschouwd. In marktonderzoek wordt van oudsher veel gebruikgemaakt van panels. Hierbij gaat het dan bijvoorbeeld om een groep gezinnen of individuen die regelmatig wordt bemonsterd of om een herhaling van precies dezelfde steekproef bij een opinieonderzoek. De verbindende factoren tussen al deze bestanden betreffen het tijdstip en de waarnemingseenheid (Reinhard et al., 2001).

Er bestaat onderscheid tussen volledige panelbestanden (balanced) en onvolledige (unbalanced) panelbestanden. Balanced wil zeggen dat voor alle waarnemingseenheden (bedrijven, enzovoort) alle variabelen op alle tijdstippen zijn gemeten en unbalanced wil zeggen dat niet alle waarnemingseenheden op alle tijdstippen zijn waargenomen. Een roterend panelbestand is een voorbeeld van een unbalanced panel. Ieder jaar wordt een deel van de respondenten vervangen door nieuwe respondenten. Het Informatienet is een voorbeeld van een roterend en dus unbalanced panelbestand.

Balanced en unbalanced databestanden hebben beiden hun voor- en nadelen. Er zijn genoeg voorbeelden van grote paneldatabestanden die met opzet unbalanced gemaakt zijn door het panel op roterende basis op te zetten (Reinhard et al., 2001, p. 14). Gebruikers van het Informatienet geven echter al geruime tijd aan dat zij graag zouden werken met langere

tijdreeksen. Om deze reden wordt de rotatie van het Informatienet in dit rapport tegen het licht gehouden.

5.2 Doelstelling

In het huidige Informatienet wordt gebruikgemaakt van rotatie van de deelnemende bedrijven. Rotatie heeft diverse voor- en nadelen. In dit rapport worden de voor- en nadelen op een rijtje gezet, de voorkeuren van de betrokken partijen worden beschreven en er wordt een advies gegeven of de rotatietermijn aan herziening toe is. Dit advies wordt geformuleerd nadat de volgende vragen zijn beantwoord:

- Wat zijn voor- en nadelen van een roterend panel en van een vaste steekproef?
- Hoe relevant zijn de geformuleerde voor- en nadelen?
- Wat zijn de randvoorwaarden waaraan de rotatie moet voldoen?

5.3 Opzet van onderzoek

Om een advies over de methode en de termijn van rotatie in het Informatienet te kunnen geven is eerst een literatuurstudie gedaan naar de voor- en nadelen van rotatie. Daarnaast zijn ook interviews gehouden waarin de wensen van verschillende gebruikers van het Informatienet zijn geïnventariseerd. Vervolgens is de geldigheid getoetst van een aantal voor- en nadelen voor het huidige Informatienet. Op basis daarvan is een afweging gemaakt van de voor- en nadelen die voor het Informatienet van belang zijn. Bij deze afweging is tevens rekening gehouden met de randvoorwaarden waaraan het Informatienet dient te voldoen.

6. Rotatiemethodieken en randvoorwaarden

6.1 Inleiding

Allereerst wordt in dit hoofdstuk een uitleg gegeven over de huidige rotatie in het Informatienet. In de analyse van mogelijke rotatiemethodieken moet rekening worden gehouden met randvoorwaarden waar het Informatienet aan moet voldoen. De randvoorwaarden worden beschreven in paragraaf 6.2. Vervolgens worden de verschillende rotatiemethodieken in paragraaf 6.3 behandeld.

6.2 Huidige opzet Informatienet

Het aantal te werven bedrijven per jaar wordt vastgesteld door het verschil te nemen tussen het aantal gewenste bedrijven en het aantal beschikbare bedrijven. Voor het aantal bij te kiezen bedrijven per stratum geldt hetzelfde. Het aantal uit voorgaande jaren beschikbare bedrijven is per stratum bekend, alsmede het aantal afgevallen bedrijven, zodoende kan het aantal bij te kiezen bedrijven per stratum bepaald worden. Voorheen werden bedrijven binnen verschillende strata ingedeeld naar kenmerken als: grootte in ege, hectareklasse, leeftijdsklasse, regio en negtype. Momenteel wordt de indeling alleen nog gemaakt op basis van de kenmerken ege-klasse en bedrijfstype. Voor de keuze worden de bedrijven in de Landbouwtelling per stratum ingedeeld. Vervolgens worden per stratum willekeurig bedrijven getrokken.

Indien bij de werving blijkt dat het bedrijf niet geschikt is, niet in staat is tot deelname of niet wil deelnemen dan wordt een volgend bedrijf binnen het stratum, volgens aselechte methode, benaderd. Indien geen bedrijf uit hetzelfde stratum gevonden kan worden, zal eerst de ege-lassenindeling losgelaten worden. Deze situatie is echter zeer theoretisch, aangezien het aantal strata zodanig verkleind is dat in praktijk altijd een bedrijf uit hetzelfde stratum te vinden is (Vrolijk et al., 2002).¹

6.3 Randvoorwaarden

De voorwaarden die aan het Informatienet worden gesteld, zijn ook direct randvoorwaarden voor de methode van rotatie. Allereerst dient rekening te worden gehouden met het budget dat voor het werven van bedrijven en het inboeken van bedrijfsgegevens beschikbaar is. Echter, het budget is in principe niet restrictief voor de rotatiemethodiek. De inzet van middelen is flexibel doordat er ook kan worden geschoven in de bedrijfstypes (het ene type kost meer tijd om uit te werken dan het andere) en met het aantal bedrijven dat vol-

¹ Afgezien van hele specifieke groepen zoals bepaalde biologische bedrijven.

gens de EU of volgens de volledige variant wordt uitgewerkt. Wel geldt dat extra middelen moeten worden ingezet als er jaarlijks veel bedrijven vervangen moeten worden, omdat het werven en opstarten (in administratie nemen) van bedrijven extra capaciteit kost.

Een tweede restrictie is dat bedrijven af kunnen vallen. Om verschillende redenen kan een bedrijf niet meer bereid zijn om mee te werken aan het Informatienet van het LEI. De redenen hiervoor zijn in de praktijk vaak moeilijk te achterhalen. Meestal komt het erop neer dat men het te druk heeft, of dat men er niet voldoende belangstelling voor heeft. Soms heeft men elders een deelboekhouding of bedrijfseconomische boekhouding opgebouwd. In een enkel geval wordt onenigheid over een geweigerde subsidie of teleurstelling in het gevoerde overheidsbeleid als reden voor weigering opgevoerd.

Een derde restrictie is het steekproefplan. Hierin wordt bepaald welke aantallen bedrijven van welke types in het Informatienet zouden moeten worden opgenomen. De verdeling van het totale aantal in administratie te nemen bedrijven over de verschillende groepen (NGE-klassen en bedrijfstypes) is vastgesteld op basis van het relatieve belang van de groep en de heterogeniteit van de groep (zie Dijk et al., 1995 en Vrolijk en Lodder, 2002). Om te kunnen voldoen aan het steekproefplan vindt jaarlijks een schoning van bedrijven plaats. Schoning kan plaatsvinden omdat bedrijven te klein (<16 ege) of te groot (>1.200 ege) zijn geworden of zijn verschoven naar een ander stratum (door bijvoorbeeld uitbreiding, inkrimping of verandering van NEG-type) dat daardoor te veel bedrijven bevat.

Een laatste restrictie is de verplichtingen die door de EU worden opgelegd aan het Informatienet. De verplichtingen houden onder andere in dat het LEI jaarlijks een boekhouding van 1.500 bedrijven dient bij te houden. Het minimaal aantal bedrijven dat in het Informatienet dient te worden opgenomen heeft effect op het minimaal aantal te werven bedrijven.

6.4 Rotatiemethodieken

In deze paragraaf wordt een beschrijving gegeven van een aantal mogelijke rotatiemethodieken. Bij een eventuele implementatie van de methodieken dient rekening gehouden te worden met de randvoorwaarden.

6.4.1 Minimale rotatie

Minimale rotatie houdt in dat er in principe geen bedrijven worden vervangen, behalve als het Informatienet niet meer aan de randvoorwaarden voldoet. De randvoorwaarde is in dit geval dat er tekorten en overschotten zijn in de verschillende strata ten opzichte van het steekproefplan. Deze tekorten en overschotten dienen te worden opgeheven en verder vindt er geen rotatie plaats. Overschotten ontstaan doordat bedrijven in een ander stratum terechtgekomen zijn, dus van bedrijfstype of van bedrijfsgrootte veranderd zijn gedurende het jaar. Ook tekorten kunnen ontstaan door verandering van stratum. Een andere mogelijkheid voor het ontstaan van tekorten is dat bedrijven hun medewerking stoppen.

6.4.2 Huidige rotatiemethode

De huidige rotatiemethode komt er op neer dat circa 25% van de bedrijven uit het Informatienet jaarlijks wordt vervangen. Ieder jaar wordt aan de hand van de meest recente Landbouwtelling die beschikbaar is, de gewenste samenstelling van de steekproef vastgesteld. Dat betekent niet dat ieder jaar een volledig nieuwe samenstelling van de steekproef wordt gekozen. De bedrijfskeuze richt zich op het aanvullen van de steekproef in verband met uitval sinds de vorige bedrijfskeuze, en het verversen van de steekproef om structuurveranderingen in de Nederlandse landbouw op te vangen. Uitgangspunt daarbij is dat bedrijven in de regel niet langer dan zes jaar in administratie worden gehouden. Omdat jaarlijks ongeveer 10% van de bedrijven die in administratie zijn, wegens uitval niet voor een volgend boekjaar beschikbaar is, zal in praktijk ieder jaar ongeveer 25% van de bedrijven vervangen worden door nieuwe bedrijven die aselekt getrokken worden (Dijk, 1989).

6.4.3 Halve rotatie

Een mogelijke alternatieve opzet is om 50% van de bedrijven vast in de steekproef te houden en 50% ieder jaar opnieuw uit het steekproefkader te trekken. In dat geval is er voor tijdreeksen en paneldata-analyses een vast aantal bedrijven beschikbaar die bovendien allemaal voor een lange periode in het Informatienet zitten. Met de huidige opzet blijven bedrijven voor zeer verschillende periodes in het boekhoudnet, variërend van één tot zeven jaar, waarbij een klein aantal zelfs langer dan zeven jaar blijft. Omdat in de alternatieve opzet elk jaar 50% van de steekproef opnieuw getrokken wordt kunnen structuurveranderingen in de populatie meegenomen worden en bovendien kan op deze wijze ook onderzocht worden of er inderdaad een leereffect optreedt en in welke mate. Het is dan namelijk mogelijk om jaarlijks de resultaten van de vaste groep bedrijven te vergelijken met die van de roterende groep bedrijven.

6.4.4 Praktische rotatie

De keuze van de rotatiemethode wordt ook beïnvloed door de aard van de schattingen die men wil maken. Naast schattingen van gemiddelden van afzonderlijke jaren, kunnen ook schattingen worden gemaakt van veranderingen tussen jaren, of schattingen op basis van meerdere jaren. Voor elk van deze schatters geldt een eigen optimale rotatiestrategie (Dijk, 1989, p. 185-189 en Cochran, 1977, p. 344-355).

- voor het schatten van de verandering in de gemiddelde waarde van een variabele tussen twee steekproeven (verschillende jaren) kan het beste de hele steekproef gehandhaafd blijven. Reden hiervoor is de variantiereductie bij gelijkblijvende steekproefelementen;
- voor het schatten van een gemiddelde over een aantal jaren samen, is het jaarlijks trekken van een nieuwe steekproef optimaal. De variantie voor deze schatter is minimaal als er geen overlap is tussen de steekproeven in opvolgende jaren;
- voor schattingen per jaar hangt het van een aantal factoren af wat de beste rotatiestrategie is. Indien gebruik gemaakt wordt van een directe schatter die alleen de gegevens van de bedrijven in het desbetreffende boekjaar gebruikt, maakt de wijze van vervanging

niet uit. Wanneer echter gebruikgemaakt wordt van een 'double sample estimate'¹ voor de bedrijven die in de steekproef blijven en de informatie uit het voorgaande jaar gebruikt wordt als hulpvariabele dan kan de variantie van deze schatter wel gereduceerd worden door de rotatiemethodiek die wordt gevolgd. De schatter zelf neemt dan de volgende vorm aan:

$$\bar{y}_2 = \left(\frac{W_{2v}}{W_{2b} + W_{2v}} \right) \bar{y}_{2v} + \left(1 - \frac{W_{2v}}{W_{2b} + W_{2v}} \right) \bar{y}_{2b}$$

Hierbij is W de inverse variantie ($1/S^2$) en de indices b en v hebben betrekking op respectievelijk blijvende en vervangen bedrijven. Voor de vervangen bedrijven is dan ook geen match te vinden en de schatter \bar{y}_{2v} is dan ook gelijk aan \bar{y}_{2v} . De schatter \bar{y}_{2b} is gelijk aan $\bar{y}_{2b} + b(\bar{y}_1 - \bar{y}_{1b})$. Hier is te zien dat ook de informatie uit het voorgaande jaar gebruikt wordt bij het maken van de schatting voor het huidige jaar. Bij een optimale vervangingsstrategie waarin gebruik wordt gemaakt van de bovenstaande schatter wordt over het algemeen meer dan 50% van de bedrijven jaarlijks vervangen. Als de correlatie tussen de vorige en de nieuwe waarneming bij bedrijven die in de steekproef blijven 0,5 is, zou het optimale vervangingspercentage 46% zijn en als deze correlatiecoëfficiënt 1 is dan zou het optimale vervangingspercentage zelfs 100% zijn.

De reductie in de varianties bij het schatten van verschillen (veranderingen tussen jaren) kunnen bij correlaties groter dan 0,7 fors zijn, met name indien weinig bedrijven vervangen worden. Cochran (p. 354) doet daarom de suggestie dat een vervanging van 1:3, 1:4 of 1:5 tussen twee opeenvolgende jaren een goede praktische regel kan zijn indien schattingen voor afzonderlijke steekproeven ('current estimators') en schattingen van verschillen tussen steekproeven beide belangrijk zijn. De praktische rotatie komt zeer dicht in de buurt van de huidige rotatie.

6.4.5 Volledige rotatie

Volledige rotatie wil zeggen dat alle bedrijven jaarlijks opnieuw worden getrokken. Hierbij kan het voorkomen dat een aantal bedrijven terugkomt in de steekproef van het Informatienet doordat zij opnieuw toevallig geselecteerd worden uit het steekproefkader. Zoals aangegeven bij de praktische rotatiemethodiek is het jaarlijks vernieuwen van de steekproef optimaal voor het schatten van een gemiddelde over een aantal jaren tegelijkertijd. Voor het schatten van veranderingen in gemiddelden is een jaarlijkse vernieuwing van de steekproefbedrijven juist niet optimaal.

¹ Een regressieschatting waarbij de hulpinformatie afkomstig is uit een andere steekproef in plaats van uit de populatie.

7. Voor- en nadelen van rotatie

7.1 Inleiding

Voor de vergelijking van de rotatiemethoden is gebruikgemaakt van een lijst met criteria. Deze criteria hebben betrekking op diverse aspecten van de methoden, zoals de mogelijkheden voor data-analyse die het Informatienet biedt indien een bepaalde rotatiemethode wordt toegepast. Ook criteria zoals de capaciteit die de verschillende rotatiemethoden vergen komen aan de orde. In dit hoofdstuk worden de verschillende rotatiemethoden beoordeeld aan de hand van de volgende criteria: bruikbaarheid voor het onderzoek, administratiekosten, leereffect, non-respons bij aanvang, uitval deelnemende bedrijven, structuurveranderingen in de landbouw, de kwaliteit van de gegevensverzameling en praktische mogelijkheden van de gegevensverzameling.

7.2 Bruikbaarheid voor het onderzoek

Voorafgaand aan de vraag welke rotatiemethode van het Informatienet leidt tot een goede bruikbaarheid voor onderzoek, wordt de vraag gesteld 'Voor welke soort onderzoek wordt het Informatienet momenteel gebruikt en waarvoor willen we het Informatienet in de toekomst gaan gebruiken?'. Deze vraag wordt beantwoord aan de hand van GDR-sessies en discussies die hebben plaatsgevonden (zie Vrolijk en Lodder (2002) en De Bont et al. (2000)).^{1 2}

Eén van de vragen die tijdens de GDR-sessie 'Doelstelling bedrijven Informatienet' aan de orde is geweest en die ook relevant is voor dit onderzoek is of men voornamelijk geïnteresseerd is in het doen van uitspraken over afzonderlijke jaren, of een groter belang stelt in ontwikkelingen over jaren. Uit de discussie kwam naar voren dat het Informatienet zowel voor onderzoek op langere termijn als voor éénjarige gemiddelden wordt gebruikt. Echter, de wens om meer lange termijngegevens te verzamelen neemt toch de overhand. Als gegevens over een lange termijn beschikbaar zijn dan kunnen lange termijnontwikkelingen in de landbouw beter gevolgd worden. Technieken als paneldata- en tijdreeksanalyse kunnen dan ook makkelijker worden toegepast. Daarnaast geldt dat lange termijngegevens ook nodig zijn om betrouwbare conclusies te kunnen trekken bij het doorrekenen van effecten. Verder geldt dat men over het algemeen minder geïnteresseerd is in het absolute niveau van bepaalde kengetal-

¹ De GDR-sessie met de naam: 'Doelstelling Bedrijven Informatienet' had tot doel een duidelijk inzicht te krijgen in de doelstellingen en eisen ten aanzien van het Informatienet. Omdat de kwaliteit van de opzet van het Informatienet alleen kan worden beoordeeld uitgaande van een verzameling wensen en eisen. Heeft plaatsgevonden op 06-07-2000.

² Studiedag LEI en Leeuwenborch (De Bont et al., 2000). In 2000 is er een studiedag georganiseerd om de samenwerking en de communicatie tussen het LEI en de Leeuwenborch op het gebied van het Informatienet te bevorderen. Zodoende kan er in de toekomst ook een betere afstemming van de activiteiten rondom het Informatienet plaatsvinden.

len, maar in de ontwikkelingen ten opzichte van andere jaren. Het LEI is ook geïnteresseerd in lange termijnontwikkelingen, maar benadrukt daarnaast veelal de jaarlijkse cijfers in haar periodieke rapportages zoals het LEB. De jaarlijkse cijfers kunnen tevens gebruikt worden om sommige korte termijn vraagstukken te beantwoorden.

Het is echter de vraag of het niet interessanter zou zijn om in de toekomst vaker gemiddelden over drie jaar te rapporteren. Dit is met name interessant als structurele zaken de nadruk verdienen. Voor gewasbescherming is het bijvoorbeeld moeilijk met jaarlijkse cijfers te werken door de grote invloed van onder andere het weer. Bij dit soort onderzoek is het dan ook beter om met driejarige cijfers te werken.

Verder is er aan de hand van de studiedag van het LEI en de Leeuwenborch de conclusie getrokken dat er grote vorderingen zijn gemaakt op het gebied van de ontwikkelingen van korte termijn aanpassingsmodellen én het gebruik van deze modellen voor beleid- en milieuonderzoek. Maar dat er nog veel onderzoek nodig is op het gebied van de dynamiek van bedrijven (onder andere investeringen en institutionele beslissingen), technologische ontwikkelingen en risico. Eén van de aanbevelingen is dan ook om de periode waarin bedrijven opgenomen zijn in het Informatienet te verlengen. Het zou veel onderzoek vergemakkelijken, onder andere op het gebied van dynamische en technologische ontwikkeling van bedrijven.

In twee van de papers die zijn opgenomen in de LEI-notitie (De Bont et al., 2000) wordt dit advies verder onderbouwd. Allereerst veronderstellen veel paneldatatechnieken dat gegevens van een bedrijf voor een groot aantal jaren beschikbaar zijn. Deze methoden komen niet tot hun recht met het huidige Informatienet. Zeer veel verschillende typen onderzoek, die zijn gebaseerd op de principes van paneldata, zouden kunnen profiteren van een lange administratieperiode: zowel onderzoek naar de dynamische ontwikkeling van bedrijven, de technische ontwikkeling, de risicoperceptie en de prijsverwachtingen maken veelvuldig gebruik van vertraagde endogene en exogene variabelen. Een langere verblijfsperiode zou de modelontwikkeling en de modeltoetsing sterk kunnen verbeteren.

Al met al zijn de paneldata- en tijdreeksmodellen en daarmee de lange termijndata van groot belang voor het onderzoek. Echter de huidige rotatiemethode van het Informatienet biedt iets te weinig mogelijkheden voor lange termijnonderzoek. De korte tijdsperiode dat bedrijven deelnemen aan het Informatienet maakt het vaak moeilijk dergelijke analysetechnieken toe te passen. De veranderingen die eventueel plaatsvinden binnen een bedrijf kunnen niet lang gevolgd worden: het gaat dan om veranderingen die op lange termijn spelen (bedrijfsovername, aankoop vaste inputs, en dergelijke). Daarnaast speelt mee dat bij een korte tijdsdimensie de asymptotische argumenten naar oneindigheid, die voor enkele toetsen wordt verondersteld (voor een uitleg wordt verwezen naar Greene, 2000) alleen vanuit de N-dimensie gerealiseerd zouden kunnen worden. Dus puur vanuit het onderzoek gezien zouden bedrijven best langer in het Informatienet mogen blijven (zo'n tien jaar) (Reinhard, 2001, p. 15). Reinhard et al. (1995) geven aan data uit het Informatienet redelijk representatief is voor structuurontwikkelingen.

Andere technische voordelen van vaste panels zijn dat zij ervoor zorgen dat individuele heterogeniteit beter onder controle wordt gehouden doordat er controle plaatsvindt van de toestand- en tijdsvariantie van variabelen die per element (bijvoorbeeld per individu) verschillen. Daarnaast geldt dat het gebruik van paneldata meer variabiliteit geeft en minder collineariteit tussen variabelen oplevert, meer vrijheidsgraden geeft en zorgt voor meer efficiency. Verder zijn onderzoekers beter in staat de dynamiek van aanpassingen te bestuderen. Ook is het beter

voor het bestuderen van de duratie van een economische toestand en de aanpassingssnelheid aan economische politiek als gebruik gemaakt kan worden van paneldata modellen (Van Asselt, 1998, p. 13).

Uit bovenstaande kan geconcludeerd worden dat in het onderzoek momenteel zowel met éénjarige gemiddelden als met lange termijnmethoden wordt gewerkt. Het zwaartepunt zal in de toekomst naar verwachting op de langere termijn komen te liggen. Mede ook wanneer de dataset van het Informatienet dit toestaat.

In de hieronder beschreven beoordeling van de verschillende rotatiemethoden is bekeken hoe zij scoren op de bruikbaarheid voor onderzoek gericht op het schatten van ontwikkelingen op de lange termijn (paneldata onderzoek), de éénjarige gemiddelden en de meerjarige gemiddelden.

Minimale rotatie:	Zeer goed bruikbaar voor het onderzoek op de lange termijn. Bedrijven worden zo lang mogelijk in de boekhouding gehouden, gegeven de randvoorwaarden. Daarnaast is deze data ook goed bruikbaar voor meerjarige gemiddelden. Voor éénjarige gemiddelden is de data ook goed bruikbaar, alleen is voor het schatten van gemiddelden over een aantal jaren heen het jaarlijks trekken van een nieuwe steekproef optimaal (Dijk, 1995, p. 186 en Cochran pp. 344-345).
Huidige rotatie:	De huidige rotatiemethode heeft een vervangingspercentage van circa 25% per jaar. Dit zorgt voor een redelijke bruikbaarheid voor paneldata doeleinden, al zou dit soort analyses gebaat zijn bij het langer volgen van bedrijven. Daarnaast is de huidige opzet bruikbaar voor zowel éénjarige gemiddelden als voor meerjarige gemiddelden. Voor het schatten van veranderingen kan de gehele steekproef het beste gehandhaafd blijven. Dit is echter niet het geval wanneer er leereffecten optreden, dan is de verandering niet representatief (Dijk, 1995, p. 186 en Cochran, 1977, pp. 344-345).
Praktische rotatie:	Praktische rotatie is bijna gelijk aan de huidige manier van rotatie. Praktische rotatie stelt namelijk een vervangingspercentage voor van 20 tot 33%. Om deze reden scoort deze methode gelijk aan de huidige rotatiemethode.
Halve rotatie:	Halve rotatie is niet echt goed bruikbaar voor lange termijnonderzoek en voor het berekenen van meerjarige gemiddelden. Dit omdat deze gegevens maar voor een klein deel van de bedrijven voorhanden zijn. Eenjarige gemiddelden kunnen wel berekend worden.
Volledige rotatie	Lange termijnonderzoek en meerjarige gemiddelden kunnen niet berekend worden. Voor éénjarige gemiddelden is deze rotatiemethode echter ideaal.

7.3 Administratiekosten

Het aanhouden van bedrijven in de steekproef is goedkoper dan elk jaar de steekproef te vernieuwen omdat er sprake is van hoge 'aanloopkosten'. Dit betekent dat het vrij veel tijd en

moeite kost om een nieuw bedrijf in de steekproef op te nemen (Van Asselt, 1998 en De Bont et al., 2000). In Reinhard (2001) wordt aangegeven dat het verhogen van de tijdsdimensie van een panel, dus het langer volgen van de deelnemers juist een kostenstijging zou inhouden. Dit argument geldt echter veelal voor consumentenonderzoek en niet voor het Informatienet. Het opnieuw benaderen van dezelfde consumenten voor een onderzoek kost relatief veel moeite, terwijl bij het Informatienet het werven van bedrijven en het inboeken van basisgegevens het eerste jaar relatief veel tijd kost. In bijlage 2 staat een berekening van de extra dagen die de werving behorende bij een bepaalde rotatiemethode kost.

Minimale rotatie:	Minste tijd nodig om bedrijven te werven.
Huidige rotatie:	Relatief weinig tijd nodig om bedrijven te werven.
Praktische rotatie:	Relatief weinig tijd nodig om bedrijven te werven.
Halve rotatie:	Relatief veel tijd nodig om bedrijven te werven.
Volledige rotatie	Meeste tijd nodig om bedrijven te werven.

7.4 Leereffect

Het leereffect wordt in alle bestaande literatuur over de rotatie als argument naar voren gebracht om de rotatietermijn kort te houden. De argumentatie over het leereffect is als volgt: het langere tijd verzamelen van gegevens en het geven van feedback op hetzelfde bedrijf zal ertoe leiden dat bedrijven zich aanpassen. Dit zal dus betekenen dat deze bedrijven niet meer een representatief beeld vormen (Vrolijk, 2000). Dijk (1989, p. 185) geeft aan dat het verwijderen van bedrijven die langer dan zes jaar in de steekproef zitten niet kan voorkomen dat er enige leereffecten bij de deelnemers aan het boekhoudnet op kunnen treden.

Boeren en tuinders die bijvoorbeeld 'eindeloos' mee zouden doen aan het Informatienet (bijvoorbeeld van 1960 tot 1990) kunnen leren van de inzichten in hun bedrijf, die zij krijgen als beloning van het meedoen aan het Informatienet. Ze kunnen daardoor zaken gaan aanpassen in het bedrijf om prestaties van het bedrijf te verhogen. Het optreden van leereffecten onder de respondenten heeft in het algemeen een negatieve invloed op de representativiteit van de steekproef (Reinhard, 2001).

Echter, het idee dat bedrijven zich zullen aanpassen komt voort uit het idee dat bedrijven over weinig informatie beschikken. Door deelname zou dan een grote hoeveelheid informatie beschikbaar komen op basis waarvan deze bedrijven hun bedrijfsvoering kunnen aanpassen. Zo'n twintig jaar geleden hadden deelnemers daadwerkelijk meer bedrijfseconomische informatie in handen dan andere landbouwbedrijven. Dit verschil is nu niet meer zo groot. Er heeft wellicht een collectieve kennistoename plaatsgevonden binnen de agrarische sector (Reinhard, 2001). Gezien de sterk toegenomen bedrijfsadministratie in de landbouw via onder andere accountants en Minas worden bedrijven veel beter op de hoogte gehouden van bedrijfseconomische aspecten van hun bedrijf (De Bont et al., 2001). In de huidige situatie beschikken de meeste bedrijven dus al over voldoende informatie (Vrolijk, 2000 en Reinhard, 2001). De kans dat de informatie een effect zal hebben op de bedrijfsvoering is in de huidige situatie dan ook veel geringer. Daarnaast is de bias ten gevolge van de non-respons waarschijnlijk aanzienlijk groter dan de bias als gevolg van een mogelijk leereffect (Vrolijk, 2000 en Reinhard, 2001).

Het eventuele leereffect voor Informatienetdeelnemers is nog nooit aangetoond. Reinhard (1999) heeft echter wel onderzoek gedaan waarbij het aantal jaren in administratie een van de verklarende variabelen voor ondernemerskwaliteit was, maar dat zegt nog weinig over de invloed van bijvoorbeeld korter of langer in administratie zijn bij het LEI. Uit het onderzoek van Reinhard (1999) bleek dat het aantal jaren dat een bedrijf deelnam aan het Informatienet wel een significant deel van de milieuefficiëntie verklaarde maar niet van de technische efficiëntie (Reinhard, 1999).

In bijlage 1 van dit rapport wordt onderzoek gedaan naar een mogelijk leereffect door van de jaren 1979, 1984, 1989, 1994 en 1999 de correlatie te onderzoeken tussen het netto-bedrijfsresultaat en het aantal jaren dat een bedrijf in boekhouding is. Op basis van deze gegevens kan in de huidige situatie geen leereffect worden aangetoond. Als er wel een leereffect zou zijn, zou dit pleiten voor een korte deelname van Informatienetbedrijven. Op dit punt scoren de verschillende rotatiemethoden als volgt:

Minimale rotatie:	Bedrijven blijven lang in boekhouding, dus het eventuele leereffect zou groot zijn.
Huidige rotatie:	Gemiddeld, bedrijven zijn vijf tot zeven jaar in boekhouding.
Praktische rotatie:	Gemiddeld, jaarlijks wordt 20 tot 33% van de bedrijven ververst.
Halve rotatie:	Gemengd, de helft van de bedrijven blijft zeer lang in boekhouding en de andere helft maar één jaar. Dus voor deze laatste groep is er sowieso geen leereffect en voor de eerste is dit eventuele effect juist heel groot.
Volledige rotatie	Bedrijven kunnen niet leren van deelname, want zij worden jaarlijks vervangen.

7.5 Non-respons bij aanvang

De non-respons van bedrijven in het Informatienet is zeer hoog. Deze schommelde in de periode 1990-2000 rond de 60%. Met een uitschieter van 90% in 1998. Deze uitschieter werd veroorzaakt doordat er bedrijven geselecteerd dienden te worden uit groepen die meestal niet mee willen werken. Dit probleem werd nog versterkt door de problemen in de varkenshouderij (Cotteleer et al., 2003).

Volgens Van Asselt (1998, p. 13) zou de non-respons toenemen met de periode waarvoor mogelijke deelnemers gevraagd worden deel te nemen. In de praktijk blijkt dit inderdaad zo te zijn. Boeren zijn niet bereid zich voor een lange periode vast te leggen. De verschillende rotatiemethoden worden als volgt beoordeeld op dit criterium:

Minimale rotatie:	Grote toename non-respons door toename deelnameperiode.
Huidige rotatie:	Toename non-respons door periode van vijf tot zeven jaar in boekhouding.
Praktische rotatie:	Toename non-respons door redelijk lange periode in boekhouding.
Halve rotatie:	Gemengd, de helft van de bedrijven blijft zeer lang in boekhouding en de andere helft maar een jaar. Dus voor deze laatste groep zal de non-respons relatief klein zijn en voor de eerste groep relatief groot.
Volledige rotatie	Geen toename non-respons bij deelname. Deelnameperiode is altijd een jaar.

7.6 Uitval deelnemende bedrijven

Bij paneldata komt uitval voor, ook wel attritie genoemd. Dit is non-respons na aanvang. Dit komt alleen voor bij paneldata omdat gegevens over meerdere perioden worden verzameld. Attritie komt door overlijden, verhuizen, of omdat men na deelname van een paar jaar geen zin meer heeft. Volgens Reinhard et al. (2001) zou de uitval toenemen met het aantal jaren dat bedrijven deelnemen aan het Informatienet. Echter, ervaring leert (bron: Jan van Dijk) dat de grootste uitval in het eerste jaar plaatsvindt en dat vervolgens het percentage uitvallers redelijk constant blijft. Uitval door overlijden, en verhuizing komt altijd voor, maar uitval doordat men geen zin meer heeft neemt juist af. Bedrijven die al jarenlang deelnemen weten op een gegeven moment niet meer beter en denken er helemaal niet meer over na om eventueel te stoppen. Deelname aan het Informatienet is voor hen een gegeven.

Ervan uitgaande dat de uitval wel toeneemt met de jaren, wordt het volgende beeld geschetst per rotatiemethode.

Minimale rotatie:	Grote toename uitval door toename deelnameperiode.
Huidige rotatie:	Toename uitval door periode van vijf tot zeven jaar in boekhouding.
Praktische rotatie:	Toename uitval door redelijk lange periode in boekhouding.
Halve rotatie:	Gemengd, de helft van de bedrijven blijft zeer lang in boekhouding en de andere helft maar één jaar. Dus voor deze laatste groep zal de uitval relatief klein zijn en voor de eerste groep relatief groot.
Volledige rotatie	Geen toename uitval bij deelname. Deelnameperiode is altijd één jaar.

7.7 Structuurveranderingen landbouw

Als bedrijven niet vervangen of ververst worden dan zouden nieuwe ontwikkelingen in de landbouw niet in het Informatienet waargenomen kunnen worden (gedurende de jaren verandert de structuur van bedrijven, worden bedrijven groter of kleiner enzovoort). Dit zou het databestand minder representatief maken voor de gehele populatie. Een roterend panel zou dus flexibeler kunnen reageren op de veranderende buitenwereld (Reinhard et al., 2001).

Praktisch geldt dat structuurveranderingen in het Informatienet wel worden meegenomen ook als de deelnameperiode van bedrijven wordt verlengd (zie Reinhard et al., 1995), zolang ondernemers gevolgd worden en niet bedrijfslocaties. Het volgen van ondernemers zorgt ervoor dat bedrijven die verhuizen naar een andere regio in het Informatienet opgenomen blijven. Als de verhuizing van een bepaald soort bedrijven naar een specifieke regio bijvoorbeeld veel voorkomt, dan worden deze structuurveranderingen voldoende weergegeven in het Informatienet. Ook groei en krimp kan worden waargenomen omdat deze ook bij deelnemende bedrijven plaatsvindt.

Minimale rotatie:	Structuurveranderingen in de landbouw worden bijna niet opgevangen in het Informatienet.
Huidige rotatie:	Structuurverandering worden redelijk opgenomen in het Informatienet.
Praktische rotatie:	Structuurverandering worden redelijk opgenomen in het Informatienet.

Halve rotatie:	Structuurverandering worden redelijk opgenomen in het Informatienet.
Volledige rotatie	Structuurveranderingen in de landbouw worden volledig opgevangen in het Informatienet.

7.8 Kwaliteit gegevensverzameling

De kwaliteit van de gegevens van bedrijven die langer meedoen, zal in het algemeen groter zijn dan van bedrijven die net starten met deelname (Dijk, 1989, pp. 185-186). Deze kwaliteit wordt onder andere verbeterd door de mogelijkheid tot terugkoppeling. Terugkoppeling betekent dat de bedrijven die deelnemen aan het Informatienet na verloop van tijd een overzicht krijgen van de uitkomsten. Daarnaast krijgen TAM's meer inzicht in de bedrijfsvoering en financiering¹ van deelnemende bedrijven.

Als we de kwaliteit van de steekproef bekijken per rotatiemethode zien we het volgende:

Minimale rotatie:	Dit zou de kwaliteit van de gegevensverzameling ten goede komen.
Huidige rotatie:	Redelijk goed voor de kwaliteit van de gegevensverzameling.
Praktische rotatie:	Redelijk goed voor de kwaliteit van de gegevensverzameling.
Halve rotatie:	Grote verschillen in kwaliteit binnen de steekproef.
Volledige rotatie	Slecht voor de kwaliteit van de steekproef.

7.9 Praktische mogelijkheden van de gegevensverzameling

Ook de praktische mogelijkheden en de gevolgen voor de gegevensverzameling spelen een rol bij de keuze van rotatiemethode. Om deze reden zijn de praktische effecten van rotatie besproken in een beperkt gezelschap waarin ook TAM's vertegenwoordigd waren. Op een zeer kwalitatieve manier zijn verbanden gelegd tussen de periode van rotatie en de effecten hiervan. In deze bijeenkomst kwamen een aantal criteria uit paragrafen 7.2 tot en met 7.8 terug. Twee mogelijke rotatievormen zijn besproken in de bijeenkomst. De eerste ging uit van de huidige methode van rotatie, die een vervanging van circa 25% van de bedrijven impliceert. De tweede is een uiterste van niet roteren. Dit behoort niet tot de mogelijke rotatiemethoden omdat niet wordt voldaan aan de randvoorwaarden die aan het Informatienet worden gesteld, maar geeft wel aanleiding tot discussie en mogelijke vernieuwende inzichten. Vervolgens zijn de resultaten gecombineerd in figuur 7.1 waar lange deelname aan het Informatienet is uitgezet ten opzichte van korte deelname.

Allereerst leidt langdurige deelname van agrarische bedrijven aan het Informatienet tot het overblijven van een selecte groep van bedrijven die voor TAM's relatief 'makkelijk' uit te werken zijn. Een dergelijke selectie geeft daardoor niet altijd een goede afspiegeling van de Nederlandse land- en tuinbouw.

¹ Er wil na verloop van tijd nog wel eens een spaarrekening bij een 2e bank opduiken.

Daartegenover staat dat de uitval naar alle waarschijnlijkheid hoger zal uitvallen als bedrijven gevraagd wordt structureel langer deel te nemen aan het Informatienet. Bij een korte deelnameperiode is de kans groot dat de periode van deelname door het LEI zelf wordt afgebroken alvorens de deelnemer dit doet. Maar ook hiervoor geldt dat de band tussen de deelnemer en de TAM de kans op uitval van een bedrijf beïnvloedt.

Ditzelfde geldt voor de respons van bedrijven in eerste instantie. Als bedrijven wordt gevraagd om voor een lange periode deel te nemen zullen zij dit eerder weigeren dan wanneer dezelfde vraag gesteld wordt voor een korte deelnameperiode. Maar ook hier geldt dat de manier waarop de TAM de deelnemer benadert van groot belang is voor de respons. Een TAM kan bijvoorbeeld aangeven dat de deelname te allen tijde beëindigd kan worden. Verder werkt het deelnemersverslag en het bedrijfsvergelijkend overzicht, dat het LEI haar deelnemers aanbiedt, nog steeds responsverhogend.

Verder zorgt het langer in boekhouding houden van bedrijven ervoor dat bedrijven minder vaak geselecteerd worden voor deelname (en geworven zouden moeten worden), terwijl ze eerder door het LEI gevraagd zijn hun deelname aan het Informatienet te beëindigen. Voor de werving geldt dat circa een op de dertig bedrijven al eerder aan het Informatienet heeft deelgenomen.

Bovendien levert werving van bepaalde bedrijfstypes en bedrijfsgroottes een zeer lage respons op. Dit geldt bijvoorbeeld voor de intensieve veehouderijbedrijven (zoals leghennenbedrijven en varkensbedrijven) kleine bedrijven en grote bedrijven. Voor grote bedrijven geldt bijvoorbeeld dat zij al veel over zichzelf weten en dat het Informatienet voor hen weinig meerwaarde heeft.

Daarnaast geldt dat informatie soms gebruikt wordt voor andere meetnetten, zoals het LMM (Landelijk Meetnet effecten Mestbeleid). Ten behoeve van dergelijke meetnetten wordt aanvullende informatie verzameld van deelnemers aan het Informatienet. Voor het LMM geldt dat deelnemers maximaal zeven jaar bemonsterd worden, maar dat de gegevens verzameling baat heeft bij de maximale deelname van zeven jaar.

Het effect van de deelnameperiode op de respons van aanvullende enquêtes en het werkplezier van TAM's is niet eenduidig. Voor aanvullende enquêtes geldt dat de respons meer afhangt van de band die er bestaat tussen deelnemers en TAM's dan van de periode dat het bedrijf al deelneemt. Een aantal langdurige deelnemers zullen het zat worden om telkens aanvullende enquêtes in te vullen, terwijl andere bedrijven altijd deel zullen nemen en daardoor een zekerheid vormen in de responderende groep. Voor het werkplezier van de TAM's kan gelden dat het uitwerken van nieuwe bedrijven een uitdaging vormen en zodoende het werkplezier vergroten. Daartegenover staan echter TAM's die het liefst hun routine aanwenden en een groter werkplezier hebben als deelnemers lang deelnemen.

De praktische gevolgen van de verschillende rotatiemethoden kunnen als volgt worden samengevat:

Minimale rotatie:	Zeker voor de korte termijn is minimale rotatie te prefereren, omdat de capaciteit van de TAM's een van de meest beperkende factoren is.
Huidige rotatie:	Was in het verleden haalbaar qua capaciteit.
Praktische rotatie:	Lijkt haalbaar qua capaciteit.
Halve rotatie:	Rotatie die per bedrijfstype verschilt zou mogelijk zijn. Dan kunnen bedrijfstypes die moeilijk te werven zijn in het Informatienet opgenomen

	blijven en dan kunnen de overige types die relatief makkelijk geworven kunnen worden geroteerd worden.
Volledige rotatie	Onhaalbaar in praktijk, zou teveel wervings- en opstartkosten met zich meebrengen.

7.10 Conclusies

Als we nu alle informatie uit de voorgaande paragrafen op een rijtje zetten en de beschreven rotatiemethode uitzetten tegen de verschillende criteria, dan krijgen we het volgende overzicht zoals te zien is in tabel 7.1.

Tabel 7.1 Voor- en nadelen van alternatieve rotatiemethoden

Rotatiemethode Criterium	Minimaal	Huidig	Praktisch	Half	Volledig
Bruikbaarheid voor onderzoek op de lange termijn (paneldata- en tijdreeksanalyse) (7.2)	++	+/-	+/-	-	--
Bruikbaarheid voor onderzoek van éénjarige gemiddelden (7.2)	+/-	+	+	+	++
Bruikbaarheid voor meerjarige gemiddelden (7.2)	+	+	+	+/-	--
Bruikbaarheid bij het inschatten van veranderingen	++	+	+	+/-	--
Benodigde capaciteit/administratiekosten (7.3)	++	+/-	+/-	-	--
Leereffect: Bedrijven zijn niet meer representatief doordat zij van deelname leren (7.4)	-	+/-	+/-	+/-	++
Uitval deelnemende bedrijven (7.5)	--	-	-	+/-	+/-
Volgen van Structuurveranderingen in de landbouw (7.6)	-	+/-	+/-	+/-	++
Non-respons bij werving (7.7)	--	+/-	+/-	+/-	++
Kwaliteit van de gegevensverzameling (7.8)	++	+/-	+/-	+/-	--
Praktische mogelijkheden gegevensverzameling (7.9)	++	+/-	+/-	+	--

Hierbij dient de opmerking gemaakt te worden dat niet alle criteria even zwaar wegen en dat de criteria en de mate waarin er effecten zijn in de loop van de tijd zijn gewijzigd. Over een aantal criteria bestaat zelfs geen eenduidig beeld. Dit zijn het leereffect en de uitval van deelnemende bedrijven. Criteria waar eenduidigheid over bestaat dienen zwaarder mee te wegen dan criteria waarvoor dit niet het geval is.

Verder is opmerkelijk dat de criteria allemaal uitersten beoordelen. Dus of de minimale vervangingstermijn wordt vereist, of de maximale. De huidige en de praktische rotatiemethoden zitten hier tussenin en worden daardoor bij de meeste criteria gemiddeld beoordeeld. Voor de halve rotatiemethode geldt dat deze moeilijk te beoordelen is op de meeste gebieden. In principe zitten beide uitersten er in, waardoor voor- en nadelen gelden voor de helft van de dataset. De hele dataset is dan ook alleen bruikbaar als er naar individuele jaren wordt gekeken.

8. Conclusies en aanbevelingen rotatieonderzoek

In de afweging tussen de verschillende methoden blijkt dat de minimale rotatie het beste gewaardeerd wordt als de beoordeling plaatsvindt aan de hand van de in hoofdstuk 6 genoemde criteria. Hierbij maakt het geen verschil of het leereffect en de uitval van deelnemende bedrijven wel of niet worden meegenomen. Minimale rotatie kan worden gezien als een natuurlijk roterend panel, waarin bedrijven worden vervangen indien zij niet meer binnen de grenzen vallen van de strata zoals aangegeven in het steekproefplan. Daarnaast is vervanging nodig doordat deelnemende bedrijven uitvallen. Verder geldt dat de optimale verdeling volgens het steekproefplan gehandhaafd blijft. Verder geldt dat de bruikbaarheid voor onderzoek prima is. Onderzoek op de lange termijn heeft hier zelfs zeer veel baat bij, terwijl onderzoek waarbij gebruik gemaakt wordt van éénjarige gemiddelden ook prima uitgevoerd kan worden. Optimaal voor deze éénjarige gemiddelden zou een volledige rotatie zijn, maar die is om capaciteitsredenen niet uitvoerbaar. Daarnaast geldt dat het actuele beeld gehandhaafd blijft en dat structuurveranderingen in de data redelijk worden meegenomen. Het actuele beeld groeit namelijk mee met de populatie door de uitval van bedrijven en de nieuwe bedrijven die hiervoor in de plaats komen, door de verschuivingen van bedrijven naar andere strata en door het steekproefplan als randvoorwaarde te stellen. De kosten die gemaakt worden voor het werven zijn relatief aan de lage kant en argumenten over het leereffect en de uitval van bedrijven zijn onduidelijke argumenten die ofwel hun geldigheid hebben verloren, ofwel tegengesproken kunnen worden. Een moeilijk punt zou de verhoogde non-respons bij het werven van bedrijven zijn als bij voorbaat wordt aangegeven dat verwacht wordt dat bedrijven oneindig lang mee doen. Dit kan natuurlijk door de TAM's ook iets minder strikt gebracht worden, waardoor de keuze aan de bedrijven wordt gelaten gedurende welke termijn zij mee zouden willen doen. Hierbij kan de wens vanuit het LEI worden uitgesproken dat bedrijven zo lang mogelijk in administratie blijven.

Het voorgaande leidt tot de aanbeveling om de methode van minimale rotatie in de toekomst in het Informatienet in te voeren.

Literatuur

Anders Ericsson, K and H.A. Simon, *Verbal Reports as Data*, *Psychological Review*. Vol. 87, no. 3, 1980, pp. 215-250.

Asselt, M. van, *De steekproef op de som, Een onderzoek van non-respons en stratificatie bij de steekproef voor het Bedrijven-Informatienet van LEI-DLO*. Afstudeerscriptie, 1998.

Bont, C.J.A.M. de, A. Oude Lansink en J. Peerlings, *Het Bedrijven-Informatienet tien jaar van onderzoek, wat zijn de volgende stappen*. Notitie 00.10. 2000.

Cochran, W.G., *Sampling techniques*. third edition. Wiley USA, 1977.

Cotteleer, G., K. Gardebroek, H.C.J. Vrolijk en W. Dol., *Opfriscursus statistiek*. Rapport 8.03.05. LEI, Den Haag, 2003.

Dijk, J., *De steekproef gewogen, een evaluatie van het LEI-boekhoudnet van landbouwbedrijven*. Onderzoekverslag 53. LEI, Den Haag, 1989.

Dijk, J., K. Lodder, J. Luyt en H.C. Pruis, *Voorstel voor de indeling van de populatie land- en tuinbouwbedrijven in groepen en bijbehorend keuzeplan*. Interne nota 422. LEI, Den Haag, 1995.

Dijk, van J.P.M., K. Lodder en H.C.J. Vrolijk, *De steekproef voor het Bedrijven-Informatienet van het LEI; Bedrijfskeuze 2000 en selectieplan 2001*. Report 1.01.02. LEI, Den Haag, 2002.

Dijk, van J.P.M., J.J.P. Groot, K. Lodder en H.C.J. Vrolijk, *De steekproef voor het Bedrijven-Informatienet van het LEI; Bedrijfskeuze 1999 en selectieplan 2000*. Rapport 6.00.94. LEI, Den Haag, 2000.

Dijk, van J.P.M., J.J.P. Groot, K. Lodder, L.C. van Staalduinen en H.C.J. Vrolijk, *De steekproef voor het Bedrijven-Informatienet van het LEI; Bedrijfskeuze 1998 en selectieplan 1999*, Rapport 6.99.94. LEI, Den Haag, 1999.

Greene, W.H., *Econometric analysis*, third edition, Prentice-Hall International Inc., 1997.

Kassarjian, H.H., *Content Analysis in Consumer Research*, *Journal of Consumer Research*, Vol. 4, pp. 8-16, 1977.

Morris, R., *Computerized content analysis in management research: a demonstration of advantages and limitations*. Journal of Management, Vol. 20, no. 4, 1994, pp. 903-931.

Poppe, K.J., *Het Bedrijven-Informatienet van A tot Z*. Rapport 1.03.06. LEI, Den Haag, 2004.

Reinhard, A.J., F. Bouma, J. Dijk en D. Oudendag, *Weergave van structurele ontwikkelingen in het LEI-boekhoudnet*. Onderzoekverslag 129. LEI, Den Haag, 1995.

Reinhard, S., *Econometric analysis of economic and environmental efficiency of Dutch dairy farms*. PhD thesis, Wageningen University, 1999.

Reinhard, S., L. van Staaldhuizen en M. Spijkerman, *Handleiding voor de mogelijkheden en het gebruik van paneldata op het LEI, Het Informatienet en de Landbouwtelling*. Notitie 01.03. LEI, Den Haag, 2001.

Vrolijk, H.C.J. en K. Lodder, *Voorstel tot vernieuwing van het steekproefplan voor het Bedrijven-Informatienet*, Rapport 1.02.02, LEI, Den Haag, 2002.

Vrolijk, H.C.J., K. Lodder en H.B. van der Veen, *Informatienet in zicht; totstandkoming en kwaliteit van de steekproef land- en tuinbouw van het Bedrijven-Informatienet 2001*. LEI, Den Haag, 2004.

Vrolijk, H.C.J., *STARS: statistics for regional studies, Proceedings of Pacioli 11; New roads for farm accounting and FADN*. Report 8.04.01. LEI, The Hague, 2004.

Vrolijk, H.C.J., G. Cotteleer, J.P.M. van Dijk en K. Lodder, *De steekproef voor het Bedrijven-Informatienet van het LEI; Bedrijfskeuze 2001 en selectieplan 2002*. Rapport 1.02.04. LEI, Den Haag, 2002.

Vrolijk, H.C.J. en W. Dol, *Statistics for regional studies: User guide*, Working Paper, LEI, Den Haag, 2003.

Vrolijk, H.C.J., W. Dol en G. Cotteleer, *Het schatten van kenmerken van kleine deelgebieden*. Rapport 8.02.05. LEI, Den Haag, 2002.

Vrolijk, H.C.J. en K. Lodder, *Voorstel tot vernieuwing van het steekproefplan voor het Bedrijven-Informatienet*. Rapport 1.02.02. LEI, Den Haag, 2002.

Vrolijk, H.C.J., 'Working Procedures for the selection of farms in the FADN', In: G. Beers en K. Poppe (eds.), *PACIOLI 9; Innovations in the FADN*. Report 8.02.02. Agricultural Economics Research Institute (LEI), The Hague, 2002.

Vrolijk, H.C.J., K. Lodder en H.B. van der Veen (2004), *Bin in zicht; totstandkoming en kwaliteit van de steekproef land- en tuinbouw van het Bedrijven-Informatienet 2001*, Rapport 1.04.01, LEI, Den Haag.

Zwart, P.S. , Methoden van Marktonderzoek, Stenfert Kroese, Amsterdam, 1994.

Bijlage 1 Samenhang Informatienet-deelname en prestaties bedrijf

De prestaties van een bedrijf worden gemeten aan de hand van het nettobedrijfsresultaat en de Informatienet-deelname is het aantal jaren dat een bedrijf al deelneemt aan het Informatienet. De hypothese wordt getoetst of bedrijven beter gaan presteren naarmate zij langer aan het Informatienet deelnemen. Om dit te toetsen wordt nagegaan of bedrijven die langer deelnemen aan het Informatienet ook een hoger nettobedrijfsresultaat hebben.

Om een eerste indruk te krijgen is het nettobedrijfsresultaat in een grafiek uitgezet tegen het aantal jaren dat bedrijven in boekhouding zijn. Vervolgens zijn de gemiddelde nettobedrijfsresultaten uitgedraaid voor groepen die hetzelfde aantal jaren in boekhouding zijn. Als de hypothese waar is dat bedrijven ervan leren als zij in boekhouding zijn bij het LEI, zouden we verwachten dat het gemiddelde nettobedrijfsresultaat zou oplopen met het aantal jaren dat een bedrijf in boekhouding is. Deze analyses leverden geen resultaat op doordat de verschillen tussen bedrijven veel groter zijn en daarmee een veel groter deel van de variantie verklaren dan het aantal jaren dat een bedrijf aan het Informatienet deelneemt. Vervolgens is gekeken naar de correlatiecoëfficiënten tussen het netto bedrijfsresultaat en het aantal jaren dat een bedrijf in boekhouding is. De correlatiecoëfficiënt heeft dezelfde waarde als de genormaliseerde beta van een regressie van het netto bedrijfsresultaat op het aantal jaren dat een bedrijf in boekhouding is. Dus een regressie-analyse geeft geen extra informatie. We verwachten een positieve (hoge) correlatie als de bovenstaande hypothese waar zou zijn. Deze analyse leverde wel positieve resultaten. In de onderstaande tabellen zijn de correlaties afgebeeld voor de afzonderlijke jaren.

1999

Methode	Coefficiënt	Significantie (2-tailed)
Pearson's correlatiecoëfficiënt	0,022	0,410
Kendall's tau b	-0,015	0,440
Spearman's rho	-0,021	0,438

1994

Methode	Coefficiënt	Significantie (2-tailed)
Pearson's correlatiecoëfficiënt	-0,014	0,567
Kendall's tau b	-0,001	0,949
Spearman's rho	-0,002	0,931

1989

Methode	Coefficiënt	Significantie (2-tailed)
Pearson's correlatiecoëfficiënt	0,116	0,000
Kendall's tau b	0,033	0,107
Spearman's rho	0,046	0,105

** Correlation is significant at the 0.01 level (2-tailed).

1984

Methode	Coefficiënt	Significantie (2-tailed)
Pearson's correlatiecoëfficiënt	-0,027	0,254
Kendall's tau b	-0,056	0,001
Spearman's rho	-0,080	0,001

1979

Methode	Coefficiënt	Significantie (2-tailed)
Pearson's correlatiecoëfficiënt	0,054	0,019
Kendall's tau b	0,029	0,076
Spearman's rho	0,040	0,081

Uit deze tabellen blijkt dat er in de jaren negentig geen positieve samenhang bestond tussen het resultaat en het aantal jaren van deelname. Deze positieve correlatie is wel zichtbaar voor de jaren 89 en 79. Deze resultaten lijken de hypothese bevestigen dat deelnemers in eerdere jaren een duidelijk voordeel hadden door deel te nemen. In recente jaren zou dit voordeel zijn afgenomen doordat bedrijven uit andere bronnen al over een grote hoeveelheid informatie beschikken.

Bijlage 2 Berekening benodigde capaciteit per methode

De calculatie van de benodigde capaciteit is gebaseerd op de kosten die gemaakt worden voor extra werving (zie ook de notitie betrouwbaarheden Informatienet). De kosten worden weergegeven in dagen die de extra werving kost. Het uitgangspunt in de notitie over betrouwbaarheden in het Informatienet is dat het werven van nieuwe bedrijven één uur kost en het in boekhouding nemen van nieuwe bedrijven twee dagen. De uitgangspunten over de afvallers worden gebaseerd op de uitvallers in de jaren 1996-2001 (Dijk et al. 1999, 2000, 2002 en Vrolijk et al., 2002).

In het verleden gold dat landbouwbedrijven sowieso afvielen als zij zeven jaar in de boekhouding waren geweest en voor tuinbouwbedrijven gold een vervangingsperiode van vijf jaar. Verder was er sprake van natuurlijk verloop. Natuurlijk verloop waren de bedrijven die verschoven naar andere strata en bedrijven die om verschillende redenen niet meer wensten deel te nemen. In de berekeningen van het percentage afvallers, indien er sprake zou zijn van minimale rotatie, wordt uitgegaan van 25% afvallers van bedrijven die nieuw worden geworven in het eerste jaar (dit zijn bedrijven waarvan vaak geen verslag uitgedraaid kan worden aan het eind van het jaar). Het percentage afvallers wordt naarmate bedrijven langer meedoen in het Informatienet kleiner (Jan van Dijk).

Met behulp van de bovenstaande gegevens wordt vervolgens het gemiddelde percentage uitvallers berekend voor de methode van minimale rotatie. Als er sprake is van minimale rotatie dan zal de uitval dus ongeveer 10% bedragen. We gaan ervan uit dat jaarlijks 150 van de 1.500 bedrijven moeten worden vervangen. Dit betekent een extra druk op de capaciteit van 319 dagen.

Het huidige vervangingsbeleid komt erop neer dat circa 25% van de 1.500 bedrijven jaarlijks wordt vervangen. Dit betekent dat er jaarlijks een extra capaciteit van circa 797 dagen nodig is ($375 \times 2,125$). Halve rotatie betekent dat 50% van de bedrijven jaarlijks sowieso vervangen wordt en dat daarnaast nog eens 10% van het vaste panel afvalt. De jaarlijkse vervanging komt dan neer op circa 825 bedrijven. Dit vergt een extra jaarlijkse capaciteit van ongeveer 1.753 dagen.

De praktische rotatie geeft een vervanging van 20 tot 33% aan. Dit betekent een extra benodigde capaciteit van 638 tot 1.052 dagen. De volledige rotatie gaat ervan uit dat de gehele steekproef wordt vervangen. In deze berekening gaan we ervan uit dat er geen bedrijven toevallig terugkomen in de steekproef. Dit betekent dat de extra benodigde capaciteit 3.188 dagen bedraagt.

Rotatiemethode	Benodigde capaciteit werving (in dagen)
Minimale rotatie	319
Huidige rotatie	797
Halve rotatie	1.753
Praktische rotatie	638 tot 1.052
Volledige rotatie	3.188